GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:56:59 ; Search time 34 Seconds
(without alignments)
3248.427 Million cell updates/sec

Title: US-10-091-085-7
Perfect score: 2252
Sequence: 1 MATSWGTVFFMLVVSCVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2221	98.6	428	4	Q96RX0	Q96rx0 homo sapien
2	2089	92.8	407	4	Q8WUB3	Q8wub3 homo sapien
3	1975.5	87.7	427	11	Q8CD29	Q8cd29 mus musculu
4	1969.5	87.5	427	11	Q8BR23	Q8br23 mus musculu
5	993	44.1	483	4	Q8TAS7	Q8tas7 homo sapien
6	993	44.1	503	4	Q8N3H3	Q8n3h3 homo sapien
7	693.5	30.8	461	5	O76268	O76268 drosophila
8	693.5	30.8	464	5	Q9VQI8	Q9vqi8 drosophila
9	616.5	27.4	479	5	Q9XU84	Q9xu84 caenorhabdi
10	590	26.2	278	11	Q8CHZ3	Q8chz3 mus musculu
11	535.5	23.8	556	3	Q9UT35	Q9ut35 schizosacch
12	513.5	22.8	489	10	Q8H7L6	Q8h7l6 oryza sativ
13	513.5	22.8	599	3	Q8TGH6	Q8tgh6 candida alb
14	510	22.6	522	3	Q9HEM6	Q9hem6 kluyveromyc
15	507.5	22.5	467	10	Q9SPM7	Q9spm7 dolichos bi
16	502.5	22.3	455	10	Q9SPM6	Q9spm6 medicago sa
17	493	21.9	462	10	Q9XFC9	Q9xfc9 dolichos bi
18	492	21.8	463	10	Q9FVC3	Q9fvc3 glycine soj
19	486.5	21.6	456	10	Q9SPM8	Q9spm8 lotus japon
20	484	21.5	455	10	Q9FEA6	Q9fea6 pisum sativ
21	484	21.5	472	10	Q8L704	Q8l704 arabidopsis
22	483	21.4	472	10	Q9SPM5	Q9spm5 arabidopsis
23	481	21.4	455	10	Q8GTB1	Q8gtb1 pisum sativ
24	480	21.3	472	10	Q9M7B3	Q9m7b3 arabidopsis
25	479	21.3	471	10	Q9SQG2	Q9sqg2 arabidopsis
26	476	21.1	447	10	Q9SLV4	Q9slv4 pisum sativ
27	471	20.9	467	10	Q8RVT6	Q8rvt6 pisum sativ
28	470	20.9	468	10	Q9FVC2	Q9fvc2 glycine soj
29	467	20.7	407	10	Q9AVN8	Q9avn8 pisum sativ
30	467	20.7	455	10	Q8RVU0	Q8rvu0 pisum sativ
31	462	20.5	447	10	Q9FUI1	Q9fui1 pisum sativ
32	457.5	20.3	558	3	Q8TGG8	Q8tgg8 aspergillus
33	457	20.3	466	10	Q9AU15	Q9au15 medicago tr
34	444	19.7	455	10	Q8RVT9	Q8rvt9 pisum sativ
35	438	19.4	473	10	Q8RVT8	Q8rvt8 pisum sativ
36	433.5	19.2	555	10	Q94EZ2	Q94ez2 arabidopsis
37	432.5	19.2	271	5	Q8IPZ6	Q8ipz6 drosophila
38	431.5	19.2	556	3	Q9C2M0	Q9c2m0 neurospora
39	430.5	19.1	555	10	O80612	O80612 arabidopsis
40	421.5	18.7	224	11	Q8CDB6	Q8cdb6 mus musculu
41	403	17.9	493	13	Q90X66	Q90x66 gallus gall
42	398	17.7	483	10	Q9XI62	Q9xi62 arabidopsis
43	385	17.1	539	11	Q8CDV7	Q8cdv7 mus musculu
44	382	17.0	510	11	Q921Q6	Q921q6 mus musculu
45	374	16.6	690	5	Q9BHV5	Q9bhv5 leishmania

ALIGNMENTS

RESULT 1
Q96RX0
ID Q96RX0 PRELIMINARY; PRT; 428 AA.
AC Q96RX0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Pcph proto-oncogene protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20173601; PubMed=10708485;
RA Recio J.A., Zambrano N., Pena, Ld, Reig J.A., Rhoads A., Rouzaut A.,
RA Notario V.;
RT "The human PCPH proto-oncogene: cDNA identification, primary
RT structure, chromosomal mapping, and expression in normal and tumor
RT cells.";
RL Mol. Carcinog. 27:229-236(2000).
DR EMBL; AF136572; AAK82950.1; -.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
SQ SEQUENCE 428 AA; 47431 MW; F2C4F7DB650A44F6 CRC64;

Query Match 98.6%; Score 2221; DB 4; Length 428;
Best Local Similarity 98.8%; Pred. No. 3.3e-183;
Matches 423; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKAAGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
```

```
                 |||||||||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
                 ||||||||
Db        421 LQSLGISH 428


RESULT 2
Q8WUB3
ID   Q8WUB3      PRELIMINARY;      PRT;   407 AA.
AC   Q8WUB3;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to ectonucleoside triphosphate diphosphohydrolase 5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Colon;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC020966; AAH20966.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
KW   Hydrolase.
SQ   SEQUENCE   407 AA;  45336 MW;  D92A5F7DC9EC9E5B CRC64;

  Query Match             92.8%;  Score 2089;  DB 4;  Length 407;
  Best Local Similarity   99.2%;  Pred. No. 7.7e-172;
  Matches  397;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
                 |||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
                 |||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400


RESULT 3
Q8CD29
ID   Q8CD29      PRELIMINARY;      PRT;   427 AA.
AC   Q8CD29;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK031581; BAC27461.1; -.
SQ   SEQUENCE   427 AA;  47101 MW;  6E3773C842B58477 CRC64;

  Query Match             87.7%;  Score 1975.5;  DB 11;  Length 427;
  Best Local Similarity   87.4%;  Pred. No. 5.2e-162;
  Matches  373;  Conservative   25;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
                 |||||| | |||:::|| | | :| |||||||:||||||||||| | ||||||||||||
Db          1 MATSWGAV-FMLIIACVGSTVFYREQQTWFEGVFLSSMCPINVSAGTFYGIMFDAGSTGT 59

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
                 ||||||||||  |||| ||||:||||||||||||||||||||| ||||||||||||||
Db         60 RIHVYTFVQKTAGQLPFLEGEIFDSVKPGLSAFVDQPKQGAETVQELLEVAKDSIPRSHW 119

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
                 ::|||||||||||||||| ||:||| ||:|||: |||||| |||||| |  ||| |||||
Db        120 ERTPVVLKATAGLRLLPEQKAQALLLEVEEIFKNSPFLVPDGSVSIMDGSYEGILAWVTV 179

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
                 |||||||||  ||||||||||||||||||||||||||||:|||||||||||||||:|||||
Db        180 NFLTGQLHGRGQETVGTLDLGGASTQITFLPQFEKTLEKTPRGYLTSFEMFNSTFKLYTH 239

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
                 |||||||||||||||||| :|||||||||||||||||||||||||||||||||:||
Db        240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGF 299

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
                 |||||||||||:|||||||:  :|||||||||| || :|||||:||||||||||||
Db        300 EPCYAEVLRVVQGKLHQPEEVRGSAFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 359

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
                 |||||| :|:|||||||||:|||||||||||||| |:|||||||||||||||||||||
Db        360 EVCDNLGSFSSGSPFLCMDLTYITALLKDGFGFADGTLLQLTKKVNNIETGWALGATFHL 419

Qy        421 LQSLGIS 427
                 |||||||
Db        420 LQSLGIT 426


RESULT 4
Q8BR23
ID   Q8BR23      PRELIMINARY;      PRT;   427 AA.
AC   Q8BR23;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
```

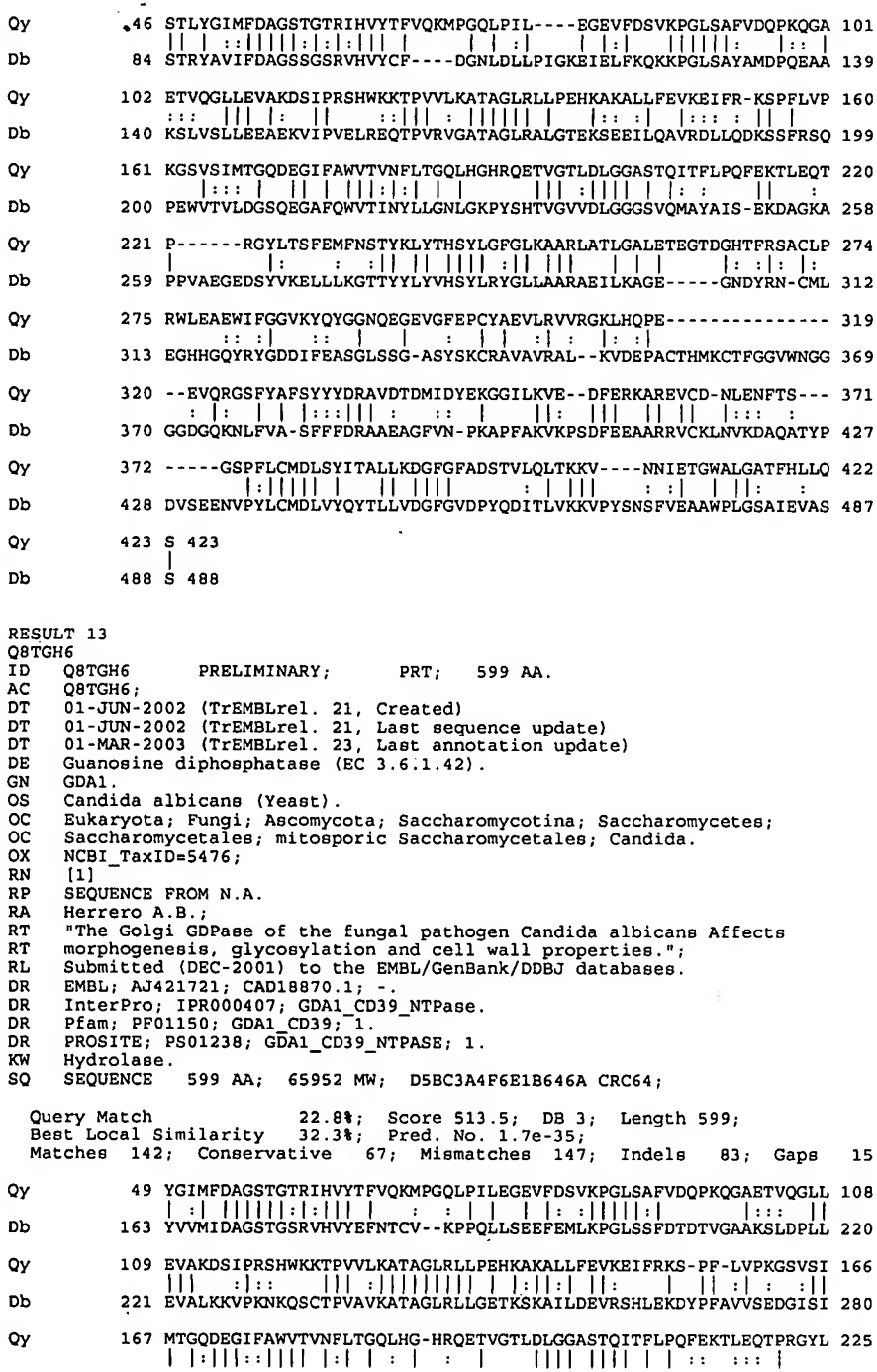
```
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK045828; BAC32507.1; -.
SQ   SEQUENCE   427 AA;  47123 MW;  87BF2CC1CCC1FCB9 CRC64;

  Query Match             87.5%;  Score 1969.5;  DB 11;  Length 427;
  Best Local Similarity   87.1%;  Pred. No. 1.7e-161;
  Matches  372;  Conservative   24;  Mismatches   30;  Indels    1;  Gaps    1;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              |||||| | |||:::|| | | :| |||||||:||||||||||| | ||||||||||||
Db          1 MATSWGAV-FMLIIACVGSTVFYREQQTWFEGVFLSSMCPINVSAGTFYGIMFDAGSTGT 59

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||  |||| ||||:||||||||||||||||||||| ||||||||||||||
Db         60 RIHVYTFVQKTAGQLPFLEGEIFDSVKPGLSAFVDQPKQGAETVQELLEVAKDSIPRSHW 119

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              ::||  |||||||||||| ||:||| ||:|||: |||||| |||||| |  ||| |||||
Db        120 ERTPXXLKATAGLRLLPEQKAQALLLEVEEIFKNSPFLVPDGSVSIMDGSYEGILAWVTV 179

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              |||||||||  ||||||||||||||||||||||||||||||||||||||||:|||||
Db        180 NFLTGQLHGRGQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTFKLYTH 239

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              |||||||||||||||||| :|||||||||||||||||||||||||||||||||:||
Db        240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGF 299

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              |||||||||:|||||||||:  :|||||||||| || :|||||||:|||||||||||
Db        300 EPCYAEVLRVVQGKLHQPEEVRGSAFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 359

Qy        361 EVCDNLENPTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              |||||| :|:||||||||||:|||||||||||||| |:|||||||||||||||||||||
Db        360 EVCDNLGSFSSGSPFLCMDLTYITALLKDGFGFADGTLLQLTKKVNNIETGWALGATFHL 419

Qy        421 LQSLGIS 427
              |||||||
Db        420 LQSLGIT 426


RESULT 5
Q8TAS7
ID   Q8TAS7      PRELIMINARY;      PRT;   483 AA.
AC   Q8TAS7;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 6 (Putative
DE   function).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC025980; AAH25980.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
KW   Hydrolase.
SQ   SEQUENCE   483 AA;  53119 MW;  A8E0E5035BCDCE8F CRC64;

  Query Match             44.1%;  Score 993;  DB 4;  Length 483;
  Best Local Similarity   52.2%;  Pred. No. 4.6e-77;
  Matches  203;  Conservative   55;  Mismatches  125;  Indels    6;  Gaps    4;

Qy         40 PINVSA---STLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQ 96
              |:  :|      |||||||||||||:||: |  : | : | |  | | ::||||||: |
Db         89 PLGTAADGHEVFYGIMFDAGSTGTRVHVFQFT-RPPRETPTLTHETFKALKPGLSAYADD 147

Qy         97 PKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSP 156
              :: |: :: ||:|||  ||   || ||:|||||||||||  ||: || :|||:|: ||
Db        148 VEKSAQGIRELLDVAKQDIPFDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASP 207

Qy        157 FLVPKGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKT 216
              |||    |||| | |||: ||:|:|||||| |       :|| ||||| |||| |||: | |
Db        208 FLVGDDCVSIMNGTDEGVSAWITINFLTGSLKTPGGSSVGMLDLGGGSTQIAFLPRVEGT 267

Qy        217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPR 275
              |: :| |||:  ||| ||||||::|||| || :|||| || :| :   ||    | ||
Db        268 LQASPPGYLTALRMFNRTYKLYSYSYLGLGLMSARLAILGGVEGQPAKDGKELVSPCLSP 327

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDR 335
              : ||    | |:  | :       | | |  |:: ::|: |||:   |||||||||
Db        328 SFKGEWEHAEVTYRVSGQKAAASLHELCAARVSEVLQNRVHRTEEVKHVDFYAFSYYYDL 387

Qy        336 AVDTDMIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395
              |    :|| |||| | | |||  |: ||  ||      ||| ||||:|:: ||:: |||
Db        388 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTLETQPQSSPFSCMDLTYVSLLLQE-FGFPR 446

Qy        396 STVLQLTKKVNNIETGWALGATFHLLQSL 424
              | ||:||:|::|:|| ||||| || : ||
Db        447 SKVLKLTRKIDNVETSWALGAIFHYIDSL 475


RESULT 6
Q8N3H3
ID   Q8N3H3      PRELIMINARY;      PRT;   503 AA.
AC   Q8N3H3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   DKFZP761J1915.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Amygdala;
RA   Ansorge W., Wirkner U., Mewes H.W., Weil B., Wiemann S.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL834158; CAD38864.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 2.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   503 AA;  54763 MW;  163933F9139D8D9F CRC64;

  Query Match             44.1%;  Score 993;  DB 4;  Length 503;
  Best Local Similarity   52.2%;  Pred. No. 4.9e-77;
  Matches  203;  Conservative   55;  Mismatches  125;  Indels    6;  Gaps    4;

Qy         40 PINVSA---STLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQ 96
              |:  :|      |||||||||||||:||: |  : | : | |  | | ::||||||: |
Db        109 PLGTAADGHEVFYGIMFDAGSTGTRVHVFQFT-RPPRETPTLTHETFKALKPGLSAYADD 167
```

```
Qy          97 PKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSP 156
               :: |: :: ||:|||  ||   || ||:|||||||||||||  ||: || :|||:|: ||
Db         168 VEKSAQGIRELLDVAKQDIPFDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASP 227

Qy         157 FLVPKGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKT 216
               |||    |||| | |||: ||:|:|||||  |       :|| ||||| |||| |||: | |
Db         228 FLVGDDCVSIMNGTDEGVSAWITINFLTGSLKTPGGSSVGMLDLGGGSTQIAFLPRVEGT 287

Qy         217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPR 275
               |: :| ||||:  ||| |||||::|||| || :|||| || :| :  ||     | ||
Db         288 LQASPPGYLTALRMFNRTYKLYSYSYLGLGLMSARLAILGGVEGQPAKDGKELVSPCLSP 347

Qy         276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDR 335
                 : ||    | |:  | :      | | | |  |:: ::|: |||:   |||||||||
Db         348 SFKGEWEHAEVTYRVSGQKAAASLHELCAARVSEVLQNRVHRTEEVKHVDFYAFSYYYDL 407

Qy         336 AVDTDMIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395
               |    :|| |||| | | |||  |: ||  ||     ||| ||||:|:: ||:: |||
Db         408 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTLETQPQSSPFSCMDLTYVSLLLQE-FGFPR 466

Qy         396 STVLQLTKKVNNIETGWALGATFHLLQSL 424
               | ||:||:|::|:|| ||||| || : ||
Db         467 SKVLKLTRKIDNVETSWALGAIFHYIDSL 495


RESULT 7
O76268
ID   O76268      PRELIMINARY;      PRT;   461 AA.
AC   O76268;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NTPASE protein (LD11641P).
GN   NTPASE OR CG3059.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98341119; PubMed=9676430;
RA   Chadwick B.P., Frischauf A.-M.;
RT   "The CD39-like gene family: identification of three new human members
RT   (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of
RT   the gene family from Drosophila melanogaster.";
RL   Genomics 50:357-367(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003581; AAF51182.1; -.
DR   EMBL; AF041048; AAC39133.1; -.
DR   EMBL; AY061134; AAL28682.1; -.
DR   FlyBase; FBgn0024947; NTPase.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
SQ   SEQUENCE   461 AA;  50845 MW;  27D00321F91A9DD1 CRC64;

  Query Match           30.8%;  Score 693.5;  DB 5;  Length 461;
  Best Local Similarity 39.1%;  Pred. No. 3.2e-51;
  Matches  156;  Conservative   73;  Mismatches  123;  Indels   47;  Gaps   12;

Qy          49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLL 108
               |  : ||||||:|:  | | :       :|  |:|   |||||:| | | :|| ::: ||
Db          79 YAAIIDAGSTGSRVLAYKFNRSFIDNKLVLYEELFKERKPGLSSFADNPAEGAHSIKLLL 138

Qy         109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMT 168
               : |:  ||: ||  ||:|||||||||||  ||: :|  |:::| || | |   :| ||
Db         139 DEARAFIPKEHWSSTPLVLKATAGLRLLPASKAENILNAVRDLFAKSEFSVDMDAVEIMD 198

Qy         169 GQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITF-------LPQFEKTLEQTP 221
               | |||||:| ||||| |:|    |      ||||| |||:||       :| ::| : :
Db         199 GTDEGIFSWFTVNFLLGRLSKTNQ--AAALDLGGGSTQVTFSPTDPDQVPVYDKYMHEV- 255

Qy         222 RGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEG--TDGHTFRSACL-PRWLE 278
                 :|| :  |    ::|||||| || ||| |      | |   :      | |: |
Db         256 ---VTSSKKIN----VFTHSYLGLGLMAARHAVF----THGYKKEDTVLESVCVNPIIAN 304

Qy         279 AEWIFGGVKYQYGGNQEGE-------VGFEPCYAEVLRVVRGKL-----HQPEEVQRGSF 326
                 | :| |:|:  | : |:       | |: |    | :|: |:       :|  ::: :
Db         305 RTWTYGNVQYKVSGKENGKSSAEQPIVDFDAC----LELVKSKVMPLVKPKPFTLKQHAV 360

Qy         327 YAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAREVC--DNLENFTSGSPFLCMDLSYIT 384
               |||||::||::: ::|   ||   || : :||:|:|   | |      ||:| ||::|:
Db         361 AAFSYYFERAIESGLVDPLAGGETTVEAYRKKAQEICAIPNDE-----QPFMCFDLTFIS 415

Qy         385 ALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHLLQS 423
               ||::|||  |   ::| ||::  |  ||||  :::| |
Db         416 TLLREGFGLNDGKKIKLYKKIDGHEISWALGCAYNVLTS 454


RESULT 8
Q9VQI8
ID   Q9VQI8      PRELIMINARY;      PRT;   464 AA.
AC   Q9VQI8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
```

```
DE   NTPASE protein.
GN   NTPASE OR CG3059.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003581; AAF51181.1; -.
DR   FlyBase; FBgn0024947; NTPase.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
SQ   SEQUENCE   464 AA;  51119 MW;  71D057ABE5AE613D CRC64;

  Query Match           30.8%;  Score 693.5;  DB 5;  Length 464;
  Best Local Similarity  39.1%;  Pred. No. 3.2e-51;
  Matches  156;  Conservative   73;  Mismatches  123;  Indels   47;  Gaps   12;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLL 108
              |  : ||||||:|:  | | :      :|  |:|   |||||:| | | :|| ::: ||
Db         82 YAAIIDAGSTGSRVLAYKFNRSFIDNKLVLYEELFKERKPGLSSFADNPAEGAHSIKLLL 141

Qy        109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMT 168
              : |:  ||: ||  ||:||||||||||||  ||: :|  |:::| || | |   :| ||
Db        142 DEARAFIPKEHWSSTPLVLKATAGLRLLPASKAENILNAVRDLFAKSEFSVDMDAVEIMD 201

Qy        169 GQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITF-------LPQFEKTLEQTP 221
              | |||||:| ||||| |:|     |      ||||| |||:||       :| ::| : :
Db        202 GTDEGIFSWFTVNFLLGRLSKTNQ--AAALDLGGGSTQVTFSPTDPDQVPVYDKYMHEV- 258

Qy        222 RGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEG--TDGHTFRSACL-PRWLE 278
                 :|| :  |    ::|||||| || ||| |      | |   :     | |: |
```

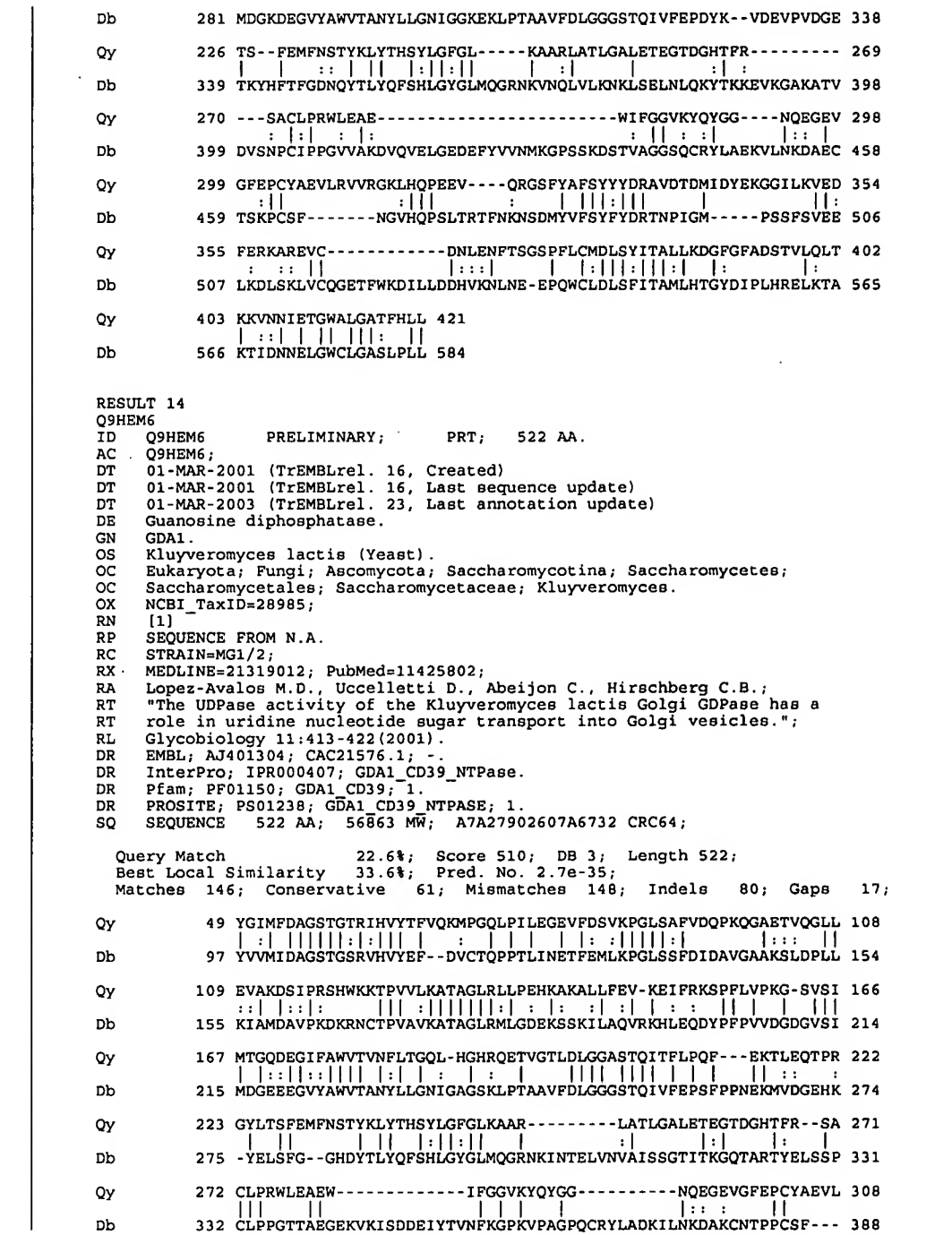
```
Db        259 ---VTSSKKIN----VFTHSYLGLGLMAARHAVF----THGYKKEDTVLESVCVNPIIAN 307

Qy        279 AEWIFGGVKYQYGGNQEGE-------VGFEPCYAEVLRVVRGKL-----HQPEEVQRGSF 326
              | :| |:|:  | : |:       | |: |    | :|: |:      :|  ::: :
Db        308 RTWTYGNVQYKVSGKENGKSSAEQPIVDFDAC----LELVKSKVMPLVKPKPFTLKQHAV 363

Qy        327 YAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAREVC--DNLENFTSGSPFLCMDLSYIT 384
              |||||::||::: ::|   ||   || : :||:|:|   | |       ||:| ||::|:
Db        364 AAFSYYFERAIESGLVDPLAGGETTVEAYRKKAQEICAIPNDE-----QPFMCFDLTFIS 418

Qy        385 ALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHLLQS 423
              ||::|||  |   ::| ||::  |  ||||  :::| |
Db        419 TLLREGFGLNDGKKIKLYKKIDGHEISWALGCAYNVLTS 457


RESULT 9
Q9XU84
ID   Q9XU84      PRELIMINARY;      PRT;   479 AA.
AC   Q9XU84;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   K08H10.4 protein.
GN   K08H10.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gardner A.E.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z83113; CAB05544.1; -.
DR   WormPep; K08H10.4; CE18877.
DR   InterPro; IPR002034; AIPM/Hcit_synth.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS00815; AIPM_HOMOCIT_SYNTH_1; 1.
SQ   SEQUENCE   479 AA;  53844 MW;  7EDC02A9D54A48ED CRC64;

  Query Match           27.4%;  Score 616.5;  DB 5;  Length 479;
  Best Local Similarity  34.1%;  Pred. No. 1.5e-44;
  Matches  149;  Conservative   82;  Mismatches  167;  Indels   39;  Gaps   16;

Qy          9 FFMLVVS--CVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYT 66
              | :|::|   : | |: : |  |  |  : :     |      : |: ||||||||:|:|
Db          6 FSILLISFFSLLSVVTTKTQY-WCHGDGVLN----NQHTCRFFTIVIDAGSTGTRLHLYK 60

Qy         67 FVQK-------MPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSH 119
              |:          || :   :| |:|  ||||||:|    |   |:::: ||: |:  :|
Db         61 FIHDPAIASHGMPFK---VEKEIFQEVKPGLSSFAKSPSSAADSLEPLLQRARKEVPHFM 117

Qy        120 WKKTPVVLKATAGLRLLPEHKAKALLFEVKE-IFRKSPFLVPKGSVSIMTGQDEGIFAWV 178
              |:||| ||||||||||    | :|  |:| ||      |      :|::| | |||:::|
Db        118 WEKTPITLKATAGLRLLPGDMADDILESVEERIFNSGFFAAFPDAVNVMPGSDEGVYSWF 177

Qy        179 TVNFLTGQLH------GHR---QETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFE 229
              |:| |    |       ||:     :|   |||| |||:|: |  |    :  ||    :
Db        178 TLNILLETLFTDEPTVGHKPAAHRSVAAFDLGGGSTQLTYWPNNEAVFSEHV-GYERDID 236

Qy        230 MFNSTYKLYTHSYLGFGLKAARLATLGALETEG--TDGHTFRSACLPRWLE-AEWIFGGV 286
              |    :|:|||:|| || ||||  |  |||:       |   ::|:|   :  || :  :
Db        237 FFGHHIRLFTHSFLGNGLIAARLNIL-QLETDNEIESTHQLITSCMPEGYQLTEWEY-AL 294

Qy        287 KYQYGGNQEGEVGFEPCYAEVLRVVR-GKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYE 345
              |: :  |      || ||      |   ::    |::    | |||::|||::: ::
```

```
Db        295 KF-WNINGSSSHSFESCYGTTKNFVESSEIMHLRELKGSPVYLFSYFFDRALNSGLVKGN 353

Qy        346 KGGILKVEDFERKAREVC----DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQL 401
Db        354 EGGKIELRQFKEAAEIACRREKTEIDDGSHWMPWQCLDLTYIYSLLRDGYQFEDNQPLVL 413

Qy        402 TKKVNNIETGWALGATF 418
Db        414 AKKIKGMEVSWGQGLAF 430


RESULT 10
Q8CHZ3
ID   Q8CHZ3      PRELIMINARY;      PRT;   278 AA.
AC   Q8CHZ3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to ectonucleoside triphosphate diphosphohydrolase 6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA   Strausberg R.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC038126; AAH38126.1; -.
KW   Hydrolase.
SQ   SEQUENCE   278 AA;  30130 MW;  75A92DD1AC76297F CRC64;

  Query Match             26.2%;  Score 590;  DB 11;  Length 278;
  Best Local Similarity   58.0%;  Pred. No. 1.3e-42;
  Matches  119;  Conservative   29;  Mismatches   51;  Indels    6;  Gaps    2;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLL 108
Db         74 YGIMFDAGSTGTRIHVFQFA-RPPGETPTLTHETFKALKPGLSAYADDVEKSAQGIQELL 132

Qy        109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMT 168
Db        133 NVAKQHIPYDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASPFLVGDDCVSIMN 192

Qy        169 GQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSF 228
Db        193 GTDEGVSAWITVNFLTGSLKTPGSSSVGMLDLGGGSTQITFLPRVEGTLQASPPGHLTAL 252

Qy        229 EMFNSTYKLYTHSYLGFGLKAARLA 253
Db        253 QMFNRTYKLYSYRWV-----CSRLA 272


RESULT 11
Q9UT35
ID   Q9UT35      PRELIMINARY;      PRT;   556 AA.
AC   Q9UT35;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative guanosine-diphosphatase (Guanosine diphosphatase).
GN   SPAC824.08 OR GDP1.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972h-;
RA   Barrell B.G., Rajandream M.A., Quail M., Seegar K., Harris D.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Sanchez R., Franco A., Notario V., Gacto M., Cansado J.;
RT   "Characterization of a guanosine diphosphatase gene from
RT   Schizosaccharomyces pombe.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL121741; CAB57338.1; -.
DR   EMBL; AF465240; AAL69974.1; -.
DR   GeneDB_SPombe; SPAC824.08; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
SQ   SEQUENCE   556 AA;  61588 MW;  1D811E3D6A6BBB85 CRC64;

  Query Match             23.8%;  Score 535.5;  DB 3;  Length 556;
  Best Local Similarity   34.1%;  Pred. No. 1.9e-37;
  Matches  143;  Conservative   61;  Mismatches  158;  Indels   57;  Gaps   13;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLL 108
Db        134 YVLMIDAGSTGSRVHVYQFNNCNPS--PKLEEEFFKMIEPGLSSFAGDPEGAAASLDPLL 191

Qy        109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKS-PFLVPKGSVSIM 167
Db        192 DYAMENVPEEYRRCSPIAVKATAGLRLTGESEAKAILKSVRQHLENDYPFPIVKDGVSIL 251

Qy        168 TGQDEGIFAWVTVNFLTGQLHGH-RQETVGTLDLGGASTQITFLPQFEKTLEQTPRG--- 223
Db        252 EGSMEGIYAWITINYLLGTLGGKATHSTVAVMDLGGASTQLVFEPRFASDGESLVDGDHK 311

Qy        224 YLTSFEMFNSTYKLYTHSYLGFGLKAARL-----------ATLGALETEGTDGHTFRSAC 272
Db        312 YVLDYN--GEQYELYQHSHLGYGLKEARKLIHKFVLNNAEALKESLELLG-DSTSIIHPC 368

Qy        273 L----------PRWLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVR----GKLHQP 318
Db        369 LHLNASLTHPDSKSEASEVVFVGPSLAHLSLQCRGIAEKALYKDKNCPVRPCSFNGVHQP 428

Qy        319 ---EEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAREVC------------ 363
Db        429 KFTETFTDSPIYLISYFYDR-----MISLGMPSTFTIEDMKYLANSVCSGPTYWQDAFSL 483

Qy        364 -DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHLL 421
Db        484 TDALKELKE-EPEWCLDLNYMISLLSVGYEIPNNRQLHTAKKIDNKELGWCLGASLSML 541


RESULT 12
Q8H7L6
ID   Q8H7L6      PRELIMINARY;      PRT;   489 AA.
AC   Q8H7L6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative apyrase.
GN   OSJNBB0014I10.10.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saski C.,
RA   Currie J., Collura K.;
RT   "Rice Genomic Sequence.";
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC126222; AAN65004.1; -.
SQ   SEQUENCE   489 AA;  52799 MW;  0BF6BF154488D38E CRC64;

  Query Match             22.8%;  Score 513.5;  DB 10;  Length 489;
  Best Local Similarity   33.7%;  Pred. No. 1.2e-35;
  Matches  142;  Conservative   70;  Mismatches  150;  Indels   59;  Gaps   16;
```



RESULT 13
Q8TGH6
ID Q8TGH6 PRELIMINARY; PRT; 599 AA.
AC Q8TGH6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Guanosine diphosphatase (EC 3.6.1.42).
GN GDA1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RA Herrero A.B.;
RT "The Golgi GDPase of the fungal pathogen Candida albicans Affects
RT morphogenesis, glycosylation and cell wall properties.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ421721; CAD18870.1; -.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW Hydrolase.
SQ SEQUENCE 599 AA; 65952 MW; D5BC3A4F6E1B646A CRC64;

Query Match 22.8%; Score 513.5; DB 3; Length 599;
Best Local Similarity 32.3%; Pred. No. 1.7e-35;
Matches 142; Conservative 67; Mismatches 147; Indels 83; Gaps 15;



RESULT 14
Q9HEM6
ID Q9HEM6 PRELIMINARY; PRT; 522 AA.
AC Q9HEM6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Guanosine diphosphatase.
GN GDA1.
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MG1/2;
RX MEDLINE=21319012; PubMed=11425802;
RA Lopez-Avalos M.D., Uccelletti D., Abeijon C., Hirschberg C.B.;
RT "The UDPase activity of the Kluyveromyces lactis Golgi GDPase has a
RT role in uridine nucleotide sugar transport into Golgi vesicles.";
RL Glycobiology 11:413-422(2001).
DR EMBL; AJ401304; CAC21576.1; -.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
SQ SEQUENCE 522 AA; 56863 MW; A7A27902607A6732 CRC64;

Query Match 22.6%; Score 510; DB 3; Length 522;
Best Local Similarity 33.6%; Pred. No. 2.7e-35;
Matches 146; Conservative 61; Mismatches 148; Indels 80; Gaps 17;

```
Qy        309 RVVRGKLHQPEEV----QRGSFYAPSYYYDRAVDTDMIDYEKGGI---LKVEDFERKARE 361
                  :|||  |    :    | |||:|||         :  |:     ::: :  ||
Db        389 ----NGIHQPSLVHTFKETSDLYVFSYFYDRT--------QPLGLPLSFTLQELQDLART 436

Qy        362 VCDNLENFTS------GS-------PFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNI 408
              ||:  | : |       ||       |  |:||::  :||  |:      |:  | : |
Db        437 VCNGEEVWESVFSGIEGSLSELSKEPQWCLDLNFQVSLLHTGYDIPLQRELRTAKTIANN 496

Qy        409 ETGWALGATFHLLQS 423
              | || |||:  ||:|
Db        497 ELGWCLGASLPLLES 511


RESULT 15
Q9SPM7
ID   Q9SPM7      PRELIMINARY;      PRT;   467 AA.
AC   Q9SPM7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Apyrase.
OS   Dolichos biflorus (Horse gram).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.
OX   NCBI_TaxID=3840;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99444909; PubMed=10517321;
RA   Roberts N.J., Brigham J., Wu B., Murphy J.B., Volpin H.,
RA   Phillips D.A., Etzler M.E.;
RT   "A Nod factor-binding lectin is a member of a distinct class of
RT   apyrases that may be unique to the legumes.";
RL   Mol. Gen. Genet. 262:261-267(1999).
DR   EMBL; AF156781; AAF00610.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
SQ   SEQUENCE   467 AA;  51164 MW;  8FCC200AA60D7376 CRC64;

  Query Match             22.5%;  Score 507.5;  DB 10;  Length 467;
  Best Local Similarity   31.5%;  Pred. No. 3.8e-35;
  Matches  130;  Conservative   72;  Mismatches  162;  Indels   49;  Gaps   12;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPIL----EGEVFDSVKPGLSAFVDQPKQGAETV 104
              | ::|||||||:|:||| | |    || :|    : |:|   ||||||: : |:: ||::
Db         57 YAVIFDAGSTGSRVHVYRFNQ----QLDLLRIGHDLELFVKTKPGLSAYAENPEEAAESL 112

Qy        105 QGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKG-S 163
                ||| |:  ||:    :||| : |||||| |    :  :|  | :: :|   |  :| :
Db        113 VPLLEEAEAVIPQELHPRTPVKVGATAGLRQLEGDASNRILQAVSDMLKKRSTLKVEGDA 172

Qy        164 VSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLE--QTP 221
              ||:::|  || : |||:|:| | |  | :||  :|||| | |: :    |   :  | |
Db        173 VSVLSGNQEGAYQWVTINYLLGNLGKHYSKTVAVVDLGGGSVQMAYAISEEDAAKAPQVP 232

Qy        222 RG---YLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLE 278
              |   |:|   :    | || |||| :|| |||   |              : |:    :
Db        233 DGVESYITEMFLRGKKYYLYVHSYLRYGLLAARAEVLKV-------SRDSENPCILSGFD 285

Qy        279 AEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVR------------GKLHQPEEVQRGS 325
                : :|||:|:      |   |  |   ||  :              |  :        :
Db        286 GYYTYGGVQYKATAPPSGS-SFSKCQNVVLEALHVNATCSYKDCTFGGIWNGGGGAGENN 344

Qy        326 FYAFSYYYDRAVDTDMID-YEKGGILKVEDFERKAREVCD-NLENFTS--------GSPF 375
              |:  |:::: | :   :|  :   |::  |||  |:  |   |::. |          |:
Db        345 FFVASFFFEVADEAGFVDPNDANAIVRPVDPEDAAKVACSTELKDLKSVFPRVKDGDVPY 404

Qy        376 LCMDLSYITALLKDGFGFADSTVLQLTKKV----NNIETGWALGATFHLLQSL 424
              :|:|| |   || ||||      : | :::    : :|  | ||:    : ||
Db        405 ICLDLVYQYTLLVDGFGIDPQQEITLVRQIQYQDSLVEAAWPLGSAIEAISSL 457
```

Search completed: January 29, 2004, 10:05:31
Job time : 35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:33:34 ; Search time 38 Seconds
(without alignments)
1787.762 Million cell updates/sec

Title: US-10-091-085-7
Perfect score: 2252
Sequence: 1 MATSWGTVFFMLVVSCVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2252	100.0	428	21	AAY44850	Human CD39-L4 vari
2	2252	100.0	428	22	AAB72240	Human CD39 like pr
3	2235	99.2	428	21	AAY44849	Human CD39-L4 prot
4	2235	99.2	428	22	AAB72238	Human CD39 like pr
5	2235	99.2	428	22	AAB72243	Human CD39 like pr
6	2235	99.2	428	23	AAE19883	Human CD39L4 prote
7	2089	92.8	405	21	AAY44851	Human CD39-L66 pro
8	2089	92.8	405	22	AAB72239	Human CD39 like pr
9	1817.5	80.7	465	23	AAE19884	Mouse CD39L4 prote
10	1645	73.0	330	21	AAB53336	Human colon cancer
11	996	44.2	456	23	AAE19881	Human CD39L2 prote
12	996	44.2	484	22	AAB72241	Human CD39 like pr
13	993	44.1	463	23	ABB06124	Human NS protein s
14	993	44.1	467	23	ABJ04657	Protein of NOVX 15
15	989	43.9	456	22	AAM93929	Human polypeptide,
16	986	43.8	446	23	ABJ04658	Protein of NOVX 15
17	894	39.7	462	22	AAU30882	Novel human secret
18	816.5	36.3	471	22	AAB72242	Mature human CD39
19	693.5	30.8	461	22	ABB66213	Drosophila melanog
20	693.5	30.8	464	22	ABB59611	Drosophila melanog
21	605	26.9	476	21	AAY70889	Protein encoded by
22	605	26.9	476	21	AAY70912	Human CD39-L4-2/3
23	548	24.3	476	21	AAY70888	Protein encoded by
24	548	24.3	476	21	AAY70911	Human CD39-L4-1 pr
25	503.5	22.4	467	20	AAW85687	DBX oligosaccharid
26	502.5	22.3	486	20	AAW85685	NBP46 root lectin.
27	502.5	22.3	486	23	AAU78819	Lotus japonicus le
28	493	21.9	462	20	AAW85684	NBP46 root lectin.
29	493	21.9	462	23	AAU78818	Dolichos biflorus
30	486.5	21.6	496	23	AAU78820	Medicago sativa le
31	484	21.5	455	22	AAB81952	Pea blight resista
32	483	21.4	472	24	ABP81286	Arabidopsis thalia
33	477.5	21.2	496	20	AAW85686	NBP46 root lectin.
34	380	16.9	495	20	AAY33296	Human membrane spa
35	371.5	16.5	458	22	AAU29271	Human PRO polypept
36	371.5	16.5	458	24	ABU71359	Human PRO5779 prot
37	371.5	16.5	458	24	ABU65816	Human secreted/tra
38	371.5	16.5	458	24	ABU66149	Novel human secret
39	371.5	16.5	458	24	ABU67653	Human secreted/tra
40	371.5	16.5	458	24	ABU65511	Human PRO polypept
41	371.5	16.5	458	24	ABU58647	Human PRO polypept
42	371.5	16.5	458	24	ABU56183	Human secreted/tra
43	371.5	16.5	458	24	ABU57178	Human PRO polypept
44	371.5	16.5	458	24	ABU10757	Human secreted/tra
45	369.5	16.4	474	21	AAY70900	Protein encoded by

ALIGNMENTS

RESULT 1
AAY44850
ID AAY44850 standard; Protein; 428 AA.
XX
AC AAY44850;
XX
DT 18-MAY-2000 (first entry)
XX
DE Human CD39-L4 variant-ACR III mutant protein.
XX
KW CD39-L4; human; apyrase; nucleotide diphosphatase; NDPase; variant;
KW ATP Diphosphohydrolase; ATPDase; adenosine diphosphate; ADP; treatment;
KW platelet aggregation; antithrombotic; thrombosis; myocardial infarction;
KW cerebral ischaemia; angina; vascular graft; extracorporeal circulation;
KW molecular weight marker; nutritional supplement; tumour; prevention;
KW drug targeting; substitution mutation.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 168
FT /note= "Wild type Asp substituted with Thr"
FT Misc-difference 170
FT /note= "Wild type Ser substituted with Gln"
FT Misc-difference 175
FT /note= "Wild type Leu substituted with Phe"
XX
PN WO200004041-A2.
XX
PD 27-JAN-2000.
XX

PF 16-JUL-1999; 99WO-US16180.
XX
PR 16-JUL-1998; 98US-0118205.
PR 24-JUL-1998; 98US-0122449.
PR 04-FEB-1999; 99US-0244444.
PR 19-MAR-1999; 99US-0273447.
PR 09-JUL-1999; 99US-0350836.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Ford J, Mulero J;
XX
DR WPI; 2000-182397/16.
DR N-PSDB; AAZ50357.
XX
PT New nucleic acid encoding human CD39-like protein, useful for treating
PT and preventing thrombotic disease -
XX
PS Claim 17; Fig 6; 125pp; English.
XX
CC The present amino acid sequence is the CD39-L4 variant, designated
CC as ACR III mutant protein, an apyrase and/or nucleotide diphosphatase
CC (NDPase). It is isolated from the human foetal liver-spleen cDNA library,
CC b2HFLS20W. It is a soluble ATP Diphosphohydrolases (ATPDase) and is
CC involved in the hydrolysis of adenosine diphosphate (ADP), the agonist
CC that causes platelet aggregation. CD39-L4 protein has 30% and 80%
CC homology to human and murine CD39. It has platelet aggregation inhibition
CC and antithrombotic activity. CD39-L4 is used to treat or prevent
CC thrombosis, myocardial infarction, cerebral ischaemia and angina. It is
CC also used in vitro, to maintain vascular grafts or during extracorporeal
CC circulation, to hydrolyse NDP, as molecular weight markers and as
CC nutritional supplements. It is used to identify therapeutic agents that
CC bind and modulate CD39-L4. It is coupled to toxins for targeting drugs
CC to tumours or other cells that express CD39-L4.
XX
SQ Sequence 428 AA;

Query Match 100.0%; Score 2252; DB 21; Length 428;
Best Local Similarity 100.0%; Pred. No. 7.4e-209;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 2
AAB72240
ID AAB72240 standard; Protein; 428 AA.
XX
AC AAB72240;
XX
DT 14-MAY-2001 (first entry)
XX
DE Human CD39 like protein CD39-L4 variant ACRIII amino acid sequence.
XX
KW Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
KW myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
KW cerebral artery thrombosis; platelet aggregation; inflammation;
KW apoptosis; autoimmune disorder; neurological disorder; mutant; mutein;
KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.
XX
OS Homo sapiens.
XX
PN WO200110205-A1.
XX
PD 15-FEB-2001.
XX
PF 09-AUG-2000; 2000WO-US21790.
XX
PR 09-AUG-1999; 99US-0370265.
PR 11-JAN-2000; 2000US-0481238.
PR 25-APR-2000; 2000US-0557800.
PR 26-MAY-2000; 2000US-0583231.
PR 30-JUN-2000; 2000US-0608285.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Ford J, Mulero JJ, Yeung G;
XX
DR WPI; 2001-147489/15.
DR N-PSDB; AAF63385.
XX
PT Polynucleotides encoding human CD39-like polypeptides, with apyrase
PT and/or NDPase activity, which are useful in the treatment of
PT pathological conditions caused by thrombosis (e.g. myocardial
PT infarction) and inflammatory disorders -
XX
PS Claim 19; Fig 6; 203pp; English.
XX
CC This invention relates to polynucleotides encoding human CD39-like
CC polypeptides with apyrase and/or NDPase activity. The polypeptides having
CC ATPDase, including NDPase, activity are useful for inhibiting platelet
CC function and can therefore be used in the prophylaxis or treatment of
CC pathological conditions caused by or involving thrombosis or excessive
CC coagulation or excessive platelet aggregation, such as myocardial
CC infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
CC artery thrombosis or intracardiac thrombosis, and conditions associated
CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
CC modulating disease states (including platelet aggregation, inflammation
CC and apoptosis) associated with ADP or other purinergic signalling by
CC reducing the levels of NDPs. The polypeptides are also useful for
CC prophylaxis or treatment of inflammation related disorders, such as
CC disorders involving sepsis or systemic inflammatory response syndrome or
CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,
CC cytokine overstimulation); autoimmune disorders such as thrombosis,
CC atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
CC neurological disorders including neurodegenerative diseases, epilepsy,
CC depression, Alzheimer's disease, Parkinson's disease, Huntington's
CC disease, and amyotrophic lateral sclerosis; and cancer. The present
CC sequence represents human CD39 like protein CD39-L4 variant ACRIII.
XX
SQ Sequence 428 AA;

Query Match 100.0%; Score 2252; DB 22; Length 428;
Best Local Similarity 100.0%; Pred. No. 7.4e-209;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 3
AAY44849
ID   AAY44849 standard; Protein; 428 AA.
XX
AC   AAY44849;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Human CD39-L4 protein.
XX
KW   CD39-L4; human; apyrase; nucleotide diphosphatase; NDPase;
KW   ATP Diphosphohydrolase; ATPDase; adenosine diphosphate; ADP; treatment;
KW   platelet aggregation; antithrombotic; thrombosis; myocardial infarction;
KW   cerebral ischaemia; angina; vascular graft; extracorporeal circulation;
KW   molecular weight marker; nutritional supplement; tumour; prevention;
KW   drug targeting; Apyrase Conserved Region; ACR.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT                   /label= Leader_peptide
FT   Protein         23..428
FT                   /label= Mature_human_CD39-L4_protein
FT                   /note= "Homologous to human and murine CD39"
FT   Binding-site    54..58
FT                   /label= ATP_Binding_region
FT   Region          129..134
FT                   /label= Apyrase_Conserved_Region
FT   Region          169..173
FT                   /label= Apyrase_Conserved_Region
FT   Region          199..206
FT                   /note= "Conserved motif in ATPDases"
XX
PN   WO200004041-A2.
XX
PD   27-JAN-2000.
XX
PF   16-JUL-1999;   99WO-US16180.
XX
PR   16-JUL-1998;   98US-0118205.
PR   24-JUL-1998;   98US-0122449.
PR   04-FEB-1999;   99US-0244444.
PR   19-MAR-1999;   99US-0273447.
PR   09-JUL-1999;   99US-0350836.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Ford J,  Mulero J;
XX
DR   WPI; 2000-182397/16.
DR   N-PSDB; AAZ50356, AAZ50359.
XX
PT   New nucleic acid encoding human CD39-like protein, useful for treating
PT   and preventing thrombotic disease -
XX
PS   Claim 15; Fig 2; 125pp; English.
XX
CC   The present amino acid sequence is the CD39-L4 protein, an apyrase
CC   and/or nucleotide diphosphatase (NDPase). It is isolated from the human
CC   foetal liver-spleen cDNA library, b2HFLS20W. It is a soluble ATP
CC   Diphosphohydrolases (ATPDase) and is involved in the hydrolysis of
CC   adenosine diphosphate (ADP), the agonist that causes platelet
CC   aggregation. CD39-L4 protein has 30% and 80% homology to human and
CC   murine CD39. It has platelet aggregation inhibition and antithrombotic
CC   activity. CD39-L4 is used to treat or prevent thrombosis, myocardial
CC   infarction, cerebral ischaemia and angina. It is also used in vitro, to
CC   maintain vascular grafts or during extracorporeal circulation, to
CC   hydrolyse NDP, as molecular weight markers and as nutritional
CC   supplements. It is used to identify therapeutic agents that bind and
CC   modulate CD39-L4. It is coupled to toxins for targeting drugs to tumours
CC   or other cells that express CD39-L4.
XX
SQ   Sequence   428 AA;

  Query Match             99.2%;  Score 2235;  DB 21;  Length 428;
  Best Local Similarity   99.3%;  Pred. No. 3.3e-207;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

```
RESULT 4
AAB72238
ID   AAB72238 standard; Protein; 428 AA.
XX
AC   AAB72238;
XX
DT   14-MAY-2001  (first entry)
XX
DE   Human CD39 like protein CD39-L4 amino acid sequence.
XX
KW   Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
KW   myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
KW   cerebral artery thrombosis; platelet aggregation; inflammation;
KW   apoptosis; autoimmune disorder; neurological disorder;
KW   Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.
XX
OS   Homo sapiens.
XX
PN   WO200110205-A1.
XX
PD   15-FEB-2001.
XX
PF   09-AUG-2000; 2000WO-US21790.
XX
PR   09-AUG-1999;   99US-0370265.
PR   11-JAN-2000; 2000US-0481238.
PR   25-APR-2000; 2000US-0557800.
PR   26-MAY-2000; 2000US-0583231.
PR   30-JUN-2000; 2000US-0608285.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Ford J,  Mulero JJ,  Yeung G;
XX
DR   WPI; 2001-147489/15.
DR   N-PSDB; AAF63383.
XX
PT   Polynucleotides encoding human CD39-like polypeptides, with apyrase
PT   and/or NDPase activity, which are useful in the treatment of
PT   pathological conditions caused by thrombosis (e.g. myocardial
PT   infarction) and inflammatory disorders -
XX
PS   Claim 15; Fig 2; 203pp; English.
XX
CC   This invention relates to polynucleotides encoding human CD39-like
CC   polypeptides with apyrase and/or NDPase activity. The polypeptides having
CC   ATPDase, including NDPase, activity are useful for inhibiting platelet
CC   function and can therefore be used in the prophylaxis or treatment of
CC   pathological conditions caused by or involving thrombosis or excessive
CC   coagulation or excessive platelet aggregation, such as myocardial
CC   infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
CC   artery thrombosis or intracardiac thrombosis, and conditions associated
CC   with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
CC   modulating disease states (including platelet aggregation, inflammation
CC   and apoptosis) associated with ADP or other purinergic signalling by
CC   reducing the levels of NDPs. The polypeptides are also useful for
CC   prophylaxis or treatment of inflammation related disorders, such as
CC   disorders involving sepsis or systemic inflammatory response syndrome or
CC   SIRS (and associated conditions such as fever, tachycardia, tachypnea,
CC   cytokine overstimulation); autoimmune disorders such as thrombosis,
CC   atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
CC   cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
CC   neurological disorders including neurodegenerative diseases, epilepsy,
CC   depression, Alzheimer's disease, Parkinson's disease, Huntington's
CC   disease, and amyotrophic lateral sclerosis; and cancer. The present
CC   sequence represents human CD39 like protein CD39-L4.
XX
SQ   Sequence   428 AA;

  Query Match            99.2%;  Score 2235;  DB 22;  Length 428;
  Best Local Similarity  99.3%;  Pred. No. 3.3e-207;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 5
AAB72243
ID   AAB72243 standard; Protein; 428 AA.
XX
AC   AAB72243;
XX
DT   14-MAY-2001  (first entry)
XX
DE   Human CD39 like protein CD39-L4 amino acid sequence.
XX
KW   Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
KW   myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
KW   cerebral artery thrombosis; platelet aggregation; inflammation;
KW   apoptosis; autoimmune disorder; neurological disorder;
KW   Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.
XX
OS   Homo sapiens.
XX
PN   WO200110205-A1.
XX
PD   15-FEB-2001.
XX
PF   09-AUG-2000; 2000WO-US21790.
XX
PR   09-AUG-1999;   99US-0370265.
PR   11-JAN-2000; 2000US-0481238.
PR   25-APR-2000; 2000US-0557800.
PR   26-MAY-2000; 2000US-0583231.
PR   30-JUN-2000; 2000US-0608285.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Ford J,  Mulero JJ,  Yeung G;
XX
DR   WPI; 2001-147489/15.
DR   N-PSDB; AAF63402.
XX
PT   Polynucleotides encoding human CD39-like polypeptides, with apyrase
PT   and/or NDPase activity, which are useful in the treatment of
PT   pathological conditions caused by thrombosis (e.g. myocardial
```

PT infarction) and inflammatory disorders -
XX
PS Example 9; Page 142-144; 203pp; English.
XX
CC This invention relates to polynucleotides encoding human CD39-like
CC polypeptides with apyrase and/or NDPase activity. The polypeptides having
CC ATPDase, including NDPase, activity are useful for inhibiting platelet
CC function and can therefore be used in the prophylaxis or treatment of
CC pathological conditions caused by or involving thrombosis or excessive
CC coagulation or excessive platelet aggregation, such as myocardial
CC infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
CC artery thrombosis or intracardiac thrombosis, and conditions associated
CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
CC modulating disease states (including platelet aggregation, inflammation
CC and apoptosis) associated with ADP or other purinergic signalling by
CC reducing the levels of NDPs. The polypeptides are also useful for
CC prophylaxis or treatment of inflammation related disorders, such as
CC disorders involving sepsis or systemic inflammatory response syndrome or
CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,
CC cytokine overstimulation); autoimmune disorders such as thrombosis,
CC atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
CC neurological disorders including neurodegenerative diseases, epilepsy,
CC depression, Alzheimer's disease, Parkinson's disease, Huntington's
CC disease, and amyotrophic lateral sclerosis; and cancer. The present
CC sequence represents the CD39 like protein CD39-L4 amino acid sequence.
XX
SQ Sequence 428 AA;

Query Match 99.2%; Score 2235; DB 22; Length 428;
Best Local Similarity 99.3%; Pred. No. 3.3e-207;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 6
AAE19883
ID AAE19883 standard; Protein; 428 AA.
XX
AC AAE19883;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human CD39L4 protein.
XX
KW Human; CD-39-like protein; CD39L4 protein; therapy; immune deficiency;
KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW rheumatoid arthritis; autoimmune thyroiditis; allergic reaction; asthma;
KW insulin dependent diabetes mellitus; periodontal disease; osteoporosis;
KW osteoarthritis; wound healing; tissue repair; Alzheimer's disease; ulcer;
KW Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
KW nervous system disease; nerve injury; ischaemia-reperfusion injury;
KW endotoxin lethality; arthritis; nephritis; inflammatory bowel disease;
KW Crohn's disease; virucide; antibacterial; antifungal; neuroprotective;
KW dermatological; immunosuppressive; vulnerary; nootropic; anticonvulsant;
KW antiinflammatory; nephrotropic; gastrointestinal; vasotropic.
XX
OS Homo sapiens.
XX
PN US6350447-B1.
XX
PD 26-FEB-2002.
XX
PF 29-JAN-1999; 99US-0240639.
XX
PR 29-JAN-1999; 99US-0240639.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Chadwick BP, Frischauf A;
XX
DR WPI; 2002-215262/27.
DR N-PSDB; AAD31695.
XX
PT An isolated polypeptide with phosphohydrolase activity, designated
PT CD39L2, useful to identify other proteins with which binding occurs or
PT identify inhibitors and for treatment of, e.g., Alzheimer's, multiple
PT sclerosis and osteoporosis -
XX
PS Example; Fig 7; 101pp; English.
XX
CC The present invention relates to novel proteins with phosphohydrolase
CC activity, designated CD-39-like (CD39L) proteins and polynucleotides
CC encoding such proteins. CD39L proteins are useful to treat infectious
CC diseases caused by viral, bacterial, fungal or other infection that may
CC be treatable with CD39L. They are useful in the treatment of various
CC immune deficiencies and disorders, autoimmune disorders such as multiple
CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune
CC thyroiditis and insulin dependent diabetes mellitus, allergic reactions
CC and conditions such as asthma and other respiratory problems, periodontal
CC disease, osteoporosis, osteoarthritis and other tooth repair processes.
CC They may have utility in compositions used for bone, cartilage, tendon,
CC ligament and/or nerve tissue growth or regeneration as well as for wound
CC healing and tissue repair and replacement and in the treatment of burns,
CC incisions and ulcers. CD39L proteins may also be useful for proliferation
CC of neural cells and for regeneration of nerve and brain tissue, i.e. for
CC the treatment of central nervous system diseases such as Alzheimer's
CC disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's
CC disease, peripheral nervous system diseases peripheral nerve injuries,
CC peripheral neuropathy and localised neuropathies. They are also used to
CC treat mechanical and traumatic disorders which involve degeneration,
CC death or trauma to neural cells or nerve tissue. CD39L proteins of the
CC invention are also useful to promote better or faster closure of non-
CC healing wounds, including pressure ulcers, ulcers associated with
CC vascular insufficiency and surgical and traumatic wounds. They also
CC exhibit anti-inflammatory activity and may be used to treat inflammatory
CC conditions including chronic or acute conditions), including ischaemia-
CC reperfusion injury, endotoxin lethality, arthritis, nephritis, cytokine
CC or chemokine-induced lung injury, inflammatory bowel disease or Crohn's
CC disease. The present sequence is human CD39L4 protein.
XX
SQ Sequence 428 AA;

Query Match 99.2%; Score 2235; DB 23; Length 428;
Best Local Similarity 99.3%; Pred. No. 3.3e-207;

```
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 7
AAY44851
ID   AAY44851 standard; Protein; 405 AA.
XX
AC   AAY44851;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Human CD39-L66 protein.
XX
KW   CD39-L4; human; CD39-L66; apyrase; nucleotide diphosphatase; NDPase;
KW   ATP Diphosphohydrolase; ATPDase; adenosine diphosphate; ADP; treatment;
KW   platelet aggregation; antithrombotic; thrombosis; myocardial infarction;
KW   cerebral ischaemia; angina; vascular graft; extracorporeal circulation;
KW   molecular weight marker; nutritional supplement; tumour; prevention;
KW   drug targeting; splice variant.
XX
OS   Homo sapiens.
XX
PN   WO200004041-A2.
XX
PD   27-JAN-2000.
XX
PF   16-JUL-1999;   99WO-US16180.
XX
PR   16-JUL-1998;   98US-0118205.
PR   24-JUL-1998;   98US-0122449.
PR   04-FEB-1999;   99US-0244444.
PR   19-MAR-1999;   99US-0273447.
PR   09-JUL-1999;   99US-0350836.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Ford J,  Mulero J;
XX
DR   WPI; 2000-182397/16.
DR   N-PSDB; AAZ50358.
XX
PT   New nucleic acid encoding human CD39-like protein, useful for treating
PT   and preventing thrombotic disease -
XX
PS   Claim 15; Page 124-125; 125pp; English.
XX
CC   The present amino acid sequence is the CD39-L66 protein, a splice
CC   variant of the CD39-L4 protein. It is an apyrase and/or nucleotide
CC   diphosphatase (NDPase), isolated from the human foetal liver-spleen cDNA
CC   library, b2HFLS20W. It is a soluble ATP Diphosphohydrolases (ATPDase) and
CC   is involved in the hydrolysis of adenosine diphosphate (ADP), the agonist
CC   that causes platelet aggregation. CD39-L4 protein has 30% and 80%
CC   homology to human and murine CD39. It has platelet aggregation inhibition
CC   and antithrombotic activity. CD39-L4 is used to treat or prevent
CC   thrombosis, myocardial infarction, cerebral ischaemia and angina. It is
CC   also used in vitro, to maintain vascular grafts or during extracorporeal
CC   circulation, to hydrolyse NDP, as molecular weight markers and as
CC   nutritional supplements. It is used to identify therapeutic agents that
CC   bind and modulate CD39-L4. It is coupled to toxins for targeting drugs
CC   to tumours or other cells that express CD39-L4.
XX
SQ   Sequence   405 AA;

  Query Match             92.8%;  Score 2089;  DB 21;  Length 405;
  Best Local Similarity   99.2%;  Pred. No. 4e-193;
  Matches  397;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
              |||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400


RESULT 8
AAB72239
ID   AAB72239 standard; Protein; 405 AA.
XX
AC   AAB72239;
XX
DT   14-MAY-2001  (first entry)
XX
DE   Human CD39 like protein CD39-L4 amino acid sequence.
XX
KW   Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
KW   myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
KW   cerebral artery thrombosis; platelet aggregation; inflammation;
KW   apoptosis; autoimmune disorder; neurological disorder;
KW   Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.
XX
OS   Homo sapiens.
XX
```

PN WO200110205-A1.
XX
PD 15-FEB-2001.
XX
PF 09-AUG-2000; 2000WO-US21790.
XX
PR 09-AUG-1999; 99US-0370265.
PR 11-JAN-2000; 2000US-0481238.
PR 25-APR-2000; 2000US-0557800.
PR 26-MAY-2000; 2000US-0583231.
PR 30-JUN-2000; 2000US-0608285.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Ford J, Mulero JJ, Yeung G;
XX
DR WPI; 2001-147489/15.
DR N-PSDB; AAF63384.
XX
PT Polynucleotides encoding human CD39-like polypeptides, with apyrase
PT and/or NDPase activity, which are useful in the treatment of
PT pathological conditions caused by thrombosis (e.g. myocardial
PT infarction) and inflammatory disorders -
XX
PS Claim 17; Page 157-158; 203pp; English.
XX
CC This invention relates to polynucleotides encoding human CD39-like
CC polypeptides with apyrase and/or NDPase activity. The polypeptides having
CC ATPDase, including NDPase, activity are useful for inhibiting platelet
CC function and can therefore be used in the prophylaxis or treatment of
CC pathological conditions caused by or involving thrombosis or excessive
CC coagulation or excessive platelet aggregation, such as myocardial
CC infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
CC artery thrombosis or intracardiac thrombosis, and conditions associated
CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
CC modulating disease states (including platelet aggregation, inflammation
CC and apoptosis) associated with ADP or other purinergic signalling by
CC reducing the levels of NDPs. The polypeptides are also useful for
CC prophylaxis or treatment of inflammation related disorders, such as
CC disorders involving sepsis or systemic inflammatory response syndrome or
CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,
CC cytokine overstimulation); autoimmune disorders such as thrombosis,
CC atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
CC neurological disorders including neurodegenerative diseases, epilepsy,
CC depression, Alzheimer's disease, Parkinson's disease, Huntington's
CC disease, and amyotrophic lateral sclerosis; and cancer. The present
CC sequence represents human CD39 like protein CD39-L4.
XX
SQ Sequence 405 AA;

Query Match 92.8%; Score 2089; DB 22; Length 405;
Best Local Similarity 99.2%; Pred. No. 4e-193;
Matches 397; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
              |||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
```

RESULT 9
AAE19884
ID AAE19884 standard; Protein; 465 AA.
XX
AC AAE19884;
XX
DT 18-JUN-2002 (first entry)
XX
DE Mouse CD39L4 protein.
XX
KW Mouse; CD-39-like protein; CD39L4 protein; therapy; immune deficiency;
KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW rheumatoid arthritis; autoimmune thyroiditis; allergic reaction; asthma;
KW insulin dependent diabetes mellitus; periodontal disease; osteoporosis;
KW osteoarthritis; wound healing; tissue repair; Alzheimer's disease; ulcer;
KW Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
KW nervous system disease; nerve injury; ischaemia-reperfusion injury;
KW endotoxin lethality; arthritis; nephritis; inflammatory bowel disease;
KW Crohn's disease; virucide; antibacterial; antifungal; neuroprotective;
KW dermatological; immunosuppressive; vulnerary; nootropic; anticonvulsant;
KW antiinflammatory; nephrotropic; gastrointestinal; vasotropic; NTPase;
KW nucleotide-triphosphatase; enzyme.
XX
OS Mus musculus.
XX
PN US6350447-B1.
XX
PD 26-FEB-2002.
XX
PF 29-JAN-1999; 99US-0240639.
XX
PR 29-JAN-1999; 99US-0240639.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Chadwick BP, Frischauf A;
XX
DR WPI; 2002-215262/27.
DR N-PSDB; AAD31696.
XX
PT An isolated polypeptide with phosphohydrolase activity, designated
PT CD39L2, useful to identify other proteins with which binding occurs or
PT identify inhibitors and for treatment of, e.g., Alzheimer's, multiple
PT sclerosis and osteoporosis -
XX
PS Example; Fig 1; 101pp; English.
XX
CC The present invention relates to novel proteins with phosphohydrolase
CC activity, designated CD-39-like (CD39L) proteins and polynucleotides
CC encoding such proteins. CD39L proteins are useful to treat infectious
CC diseases caused by viral, bacterial, fungal or other infection that may
CC be treatable with CD39L. They are useful in the treatment of various
CC immune deficiencies and disorders, autoimmune disorders such as multiple
CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune
CC thyroiditis and insulin dependent diabetes mellitus, allergic reactions
CC and conditions such as asthma and other respiratory problems, periodontal
CC disease, osteoporosis, osteoarthritis and other tooth repair processes.
CC They may have utility in compositions used for bone, cartilage, tendon,
CC ligament and/or nerve tissue growth or regeneration as well as for wound
CC healing and tissue repair and replacement and in the treatment of burns,
CC incisions and ulcers. CD39L proteins may also be useful for proliferation
CC of neural cells and for regeneration of nerve and brain tissue, i.e. for

```
CC   the treatment of central nervous system diseases such as Alzheimer's
CC   disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's
CC   disease, peripheral nervous system diseases peripheral nerve injuries,
CC   peripheral neuropathy and localised neuropathies. They are also used to
CC   treat mechanical and traumatic disorders which involve degeneration,
CC   death or trauma to neural cells or nerve tissue. CD39L proteins of the
CC   invention are also useful to promote better or faster closure of non-
CC   healing wounds, including pressure ulcers, ulcers associated with
CC   vascular insufficiency and surgical and traumatic wounds. They also
CC   exhibit anti-inflammatory activity and may be used to treat inflammatory
CC   conditions including chronic or acute conditions), including ischaemia-
CC   reperfusion injury, endotoxin lethality, arthritis, nephritis, cytokine
CC   or chemokine-induced lung injury, inflammatory bowel disease or Crohn's
CC   disease. The present sequence is mouse CD39L4 protein, also known as
CC   nucleotide-triphosphatase (NTPase).
XX
SQ   Sequence   465 AA;

  Query Match             80.7%;  Score 1817.5;  DB 23;  Length 465;
  Best Local Similarity   83.1%;  Pred. No. 9.1e-167;
  Matches  349;  Conservative   25;  Mismatches   43;  Indels    3;  Gaps    3;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              |||||| | |||:::|| | | :| |||||||:|||||||||||| | ||||||||||
Db          1 MATSWGAV-FMLIIACVGSTVFYREQQTWFEGVFLSSMCPINVSAGTFYGIMFDAGSTGA 59

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              |||||||||  |||| ||||:||||||||||||||||||||| ||||||||||||||
Db         60 RIHVYTFVQKTAGQLPFLEGEIFDSVKPGLSAFVDQPKQGAETVQELLEVAKDSIPRSHW 119

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              ::|||||||||||||| ||:||| ||:|||: |||||| |||||| |  ||| |||||
Db        120 ERTPVVLKATAGLRLLPEQKAQALLLEVEEIFKNSPFLVPDGSVSIMDGSYEGILAWVTV 179

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              |||||||| |||||||||||||||||||||||||||||||||||||:|||||
Db        180 NFLTGQLHGHRGQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTFKLYTH 239

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              |||||||||||||||||| :|||||||||||||||||||||||||||||||||:||
Db        240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGF 299

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||:|||||||||:  :|||||||||| || :|||||||:||||||||||
Db        300 EPCYAEVLRVVQGKLHQPEEVRGSAFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 359

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              |||||| :|:|||||||||:|||||||| |||:   |   |: :     | ||   ||
Db        360 EVCDNLGSFSSGSPFLCMDLTYITALLKDGLGFAERHPLTAHKE-SEQHRDW-LGLGGHL 417


RESULT 10
AAB53336
ID   AAB53336 standard; Protein; 330 AA.
XX
AC   AAB53336;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human colon cancer antigen protein sequence SEQ ID NO:876.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW   immunomodulatory; muscular; gynaecological; gastrointestinal;
KW   nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW   neural disorder; immune system disorder; muscular disorder;
KW   reproductive disorder; gastrointestinal disorder; renal disorder;
KW   infectious disease; cardiovascular disorder.
XX
OS   Homo sapiens.
XX
PN   WO200055351-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05883.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587534/55.
DR   N-PSDB; AAC98093.
XX
PT   Colon cancer associated gene sequences, referred to as colon cancer
PT   antigens, useful for the treatment, prevention, and diagnosis of colon
PT   disorders such as colon cancer -
XX
PS   Claim 11; Page 1429-1431; 2104pp; English.
XX
CC   AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC   called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC   human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC   neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC   vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC   can be used in gene therapy. The colon cancer antigen polynucleotides,
CC   proteins and antibodies to the proteins are useful for the prevention,
CC   treatment and diagnosis of colon disorders, such as colon cancer. The
CC   polynucleotides may be used in diagnostics and research, such as for
CC   chromosome identification, and as hybridisation probes. The proteins
CC   may also be used to prevent diseases such as neural disorders, immune
CC   system disorders, muscular disorders, reproductive disorders,
CC   gastrointestinal disorders, wounds, renal disorders, infectious
CC   diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
CC   AAB54007 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   330 AA;

  Query Match             73.0%;  Score 1645;  DB 21;  Length 330;
  Best Local Similarity   96.3%;  Pred. No. 2.6e-150;
  Matches  316;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        101 AETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVP 160
              | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ARAVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVP 62

Qy        161 KGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQT 220
              ||||||| | |||| |||||||||||||||||| |||||||| ||||||||||||||||
Db         63 KGSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETXGTLDLGGAXTQITFLPQFEKTLEQT 122

Qy        221 PRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAE 280
              | |||||||||||| |||||||||||||||||||||||||||||||| ||| ||||
Db        123 PXGYLTSFEMFNSTYXLYTHSYLGFGLKAARLATLGALETEGTDGHTFRSAXLPRXLEAE 182

Qy        281 WIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTD 340
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 WIFGGVKYQYGXNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTD 242

Qy        341 MIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 MIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 302

Qy        401 LTKKVNNIETGWALGATFHLLQSLGISH 428
              ||||||||||||||||||||||||||||
Db        303 LTKKVNNIETGWALGATFHLLQSLGISH 330


RESULT 11
AAE19881
ID   AAE19881 standard; Protein; 456 AA.
XX
AC   AAE19881;
XX
```

DT 18-JUN-2002 (first entry)
XX
DE Human CD39L2 protein.
XX
KW Human; CD-39-like protein; CD39L2 protein; therapy; immune deficiency;
KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW rheumatoid arthritis; autoimmune thyroiditis; allergic reaction; asthma;
KW insulin dependent diabetes mellitus; periodontal disease; osteoporosis;
KW osteoarthritis; wound healing; tissue repair; Alzheimer's disease; ulcer;
KW Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
KW nervous system disease; nerve injury; ischaemia-reperfusion injury;
KW endotoxin lethality; arthritis; nephritis; inflammatory bowel disease;
KW Crohn's disease; virucide; antibacterial; antifungal; neuroprotective;
KW dermatological; immunosuppressive; vulnerary; nootropic; anticonvulsant;
KW antiinflammatory; nephrotropic; gastrointestinal; vasotropic.
XX
OS Homo sapiens.
XX
PN US6350447-B1.
XX
PD 26-FEB-2002.
XX
PF 29-JAN-1999; 99US-0240639.
XX
PR 29-JAN-1999; 99US-0240639.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Chadwick BP, Frischauf A;
XX
DR WPI; 2002-215262/27.
DR N-PSDB; AAD31693.
XX
PT An isolated polypeptide with phosphohydrolase activity, designated
PT CD39L2, useful to identify other proteins with which binding occurs or
PT identify inhibitors and for treatment of, e.g., Alzheimer's, multiple
PT sclerosis and osteoporosis -
XX
PS Claim 1; Fig 4; 101pp; English.
XX
CC The present invention relates to novel proteins with phosphohydrolase
CC activity, designated CD-39-like (CD39L) proteins and polynucleotides
CC encoding such proteins. CD39L proteins are useful to treat infectious
CC diseases caused by viral, bacterial, fungal or other infection that may
CC be treatable with CD39L. They are useful in the treatment of various
CC immune deficiencies and disorders, autoimmune disorders such as multiple
CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune
CC thyroiditis and insulin dependent diabetes mellitus, allergic reactions
CC and conditions such as asthma and other respiratory problems, periodontal
CC disease, osteoporosis, osteoarthritis and other tooth repair processes.
CC They may have utility in compositions used for bone, cartilage, tendon,
CC ligament and/or nerve tissue growth or regeneration as well as for wound
CC healing and tissue repair and replacement and in the treatment of burns,
CC incisions and ulcers. CD39L proteins may also be useful for proliferation
CC of neural cells and for regeneration of nerve and brain tissue, i.e. for
CC the treatment of central nervous system diseases such as Alzheimer's
CC disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's
CC disease, peripheral nervous system diseases peripheral nerve injuries,
CC peripheral neuropathy and localised neuropathies. They are also used to
CC treat mechanical and traumatic disorders which involve degeneration,
CC death or trauma to neural cells or nerve tissue. CD39L proteins of the
CC invention are also useful to promote better or faster closure of non-
CC healing wounds, including pressure ulcers, ulcers associated with
CC vascular insufficiency and surgical and traumatic wounds. They also
CC exhibit anti-inflammatory activity and may be used to treat inflammatory
CC conditions including chronic or acute conditions), including ischaemia-
CC reperfusion injury, endotoxin lethality, arthritis, nephritis, cytokine
CC or chemokine-induced lung injury, inflammatory bowel disease or Crohn's
CC disease. The present sequence is human CD39L2 protein.
XX
SQ Sequence 456 AA;

Query Match 44.2%; Score 996; DB 23; Length 456;
Best Local Similarity 52.4%; Pred. No. 2.6e-87;
Matches 204; Conservative 54; Mismatches 125; Indels 6; Gaps 4;

```
Qy         40 PINVSA---STLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQ 96
              |: :|      |||||||||||||:||: |  : | : | |  | | :|||||||: |
Db         62 PLGTAADGHEVFYGIMFDAGSTGTRVHVFQFT-RPPRETPTLTHETFKAVKPGLSAYADD 120

Qy         97 PKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSP 156
               :: |: :: ||:|||  ||   || ||:||||||||||||  ||: || :|||:|: ||
Db        121 VEKSAQGIRELLDVAKQDIPFDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASP 180

Qy        157 FLVPKGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKT 216
              |||    |||| | |||: ||:|:|||| |       :|| ||||| |||| |||: | |
Db        181 FLVGDDCVSIMNGTDEGVSAWITINFLTGSLKTPGGSSVGMLDLGGGSTQIAFLPRVEGT 240

Qy        217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPR 275
              |: :| ||||:  ||| |||||::|||| || :|||| || :| :   ||    | ||
Db        241 LQASPPGYLTALRMFNRTYKLYSYSYLGLGLMSARLAILGGVEGQPAKDGKELVSPCLSP 300

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDR 335
               : ||    | |:  | :     | | | |  |:: ::|: |||:   |||||||||
Db        301 SFKGEWEHAEVTYRVSGQKAAASLHELCAARVSEVLQNRVHRTEEVKHVDFYAFSYYYDL 360

Qy        336 AVDTDMIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395
              |    :|| |||| | | |||  |: ||  ||     ||| ||||:|:: ||:: |||
Db        361 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTLETQPQSSPFSCMDLTYVSLLLQE-FGFPR 419

Qy        396 STVLQLTKKVNNIETGWALGATFHLLQSL 424
              | ||:||:|::|:|| ||||| || : ||
Db        420 SKVLKLTRKIDNVETSWALGAIFHYIDSL 448
```

RESULT 12
AAB72241
ID AAB72241 standard; Protein; 484 AA.
XX
AC AAB72241;
XX
DT 14-MAY-2001 (first entry)
XX
DE Human CD39 like protein CD39-L2 amino acid sequence.
XX
KW Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
KW myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
KW cerebral artery thrombosis; platelet aggregation; inflammation;
KW apoptosis; autoimmune disorder; neurological disorder;
KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L2.
XX
OS Homo sapiens.
XX
PN WO200110205-A1.
XX
PD 15-FEB-2001.
XX
PF 09-AUG-2000; 2000WO-US21790.
XX
PR 09-AUG-1999; 99US-0370265.
PR 11-JAN-2000; 2000US-0481238.
PR 25-APR-2000; 2000US-0557800.
PR 26-MAY-2000; 2000US-0583231.
PR 30-JUN-2000; 2000US-0608285.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Ford J, Mulero JJ, Yeung G;
XX
DR WPI; 2001-147489/15.
DR N-PSDB; AAF63386.
XX
PT Polynucleotides encoding human CD39-like polypeptides, with apyrase
PT and/or NDPase activity, which are useful in the treatment of
PT pathological conditions caused by thrombosis (e.g. myocardial
PT infarction) and inflammatory disorders -

```
XX
PS   Claim 39; Page 162-164; 203pp; English.
XX
CC   This invention relates to polynucleotides encoding human CD39-like
CC   polypeptides with apyrase and/or NDPase activity. The polypeptides having
CC   ATPDase, including NDPase, activity are useful for inhibiting platelet
CC   function and can therefore be used in the prophylaxis or treatment of
CC   pathological conditions caused by or involving thrombosis or excessive
CC   coagulation or excessive platelet aggregation, such as myocardial
CC   infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
CC   artery thrombosis or intracardiac thrombosis, and conditions associated
CC   with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
CC   modulating disease states (including platelet aggregation, inflammation
CC   and apoptosis) associated with ADP or other purinergic signalling by
CC   reducing the levels of NDPs. The polypeptides are also useful for
CC   prophylaxis or treatment of inflammation related disorders, such as
CC   disorders involving sepsis or systemic inflammatory response syndrome or
CC   SIRS (and associated conditions such as fever, tachycardia, tachypnea,
CC   cytokine overstimulation); autoimmune disorders such as thrombosis,
CC   atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
CC   cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
CC   neurological disorders including neurodegenerative diseases, epilepsy,
CC   depression, Alzheimer's disease, Parkinson's disease, Huntington's
CC   disease, and amyotrophic lateral sclerosis; and cancer. The present
CC   sequence represents human CD39 like protein CD39-L2.
XX
SQ   Sequence   484 AA;

  Query Match             44.2%;  Score 996;  DB 22;  Length 484;
  Best Local Similarity   52.4%;  Pred. No. 2.8e-87;
  Matches  204;  Conservative   54;  Mismatches  125;  Indels    6;  Gaps    4;

Qy         40 PINVSA---STLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQ 96
              |:  :|      ||||||||||||:||: |  : | : | |  | | :|||||||: |
Db         90 PLGTAADGHEVFYGIMFDAGSTGTRVHVFQFT-RPPRETPTLTHETFKAVKPGLSAYADD 148

Qy         97 PKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSP 156
              :: |: :: ||:|||  ||   || ||:|||||||||||||  ||: || :|||:|: ||
Db        149 VEKSAQGIRELLDVAKQDIPFDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASP 208

Qy        157 FLVPKGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKT 216
              |||   |||| | |||: ||:|:|||||  |      :|| |||||  |||| |||: | |
Db        209 FLVGDDCVSIMNGTDEGVSAWITINFLTGSLKTPGGSSVGMLDLGGGSTQIAFLPRVEGT 268

Qy        217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPR 275
              |: :| ||||:  ||| |||||::|||| || :|||| || :| :   ||    | ||
Db        269 LQASPPGYLTALRMFNRTYKLYSYSYLGLGLMSARLAILGGVEGQPAKDGKELVSPCLSP 328

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDR 335
              -: ||    | |:  | :       | | | |  |:: ::|: |||:   |||||||||
Db        329 SFKGEWEHAEVTYRVSGQKAAASLHELCAARVSEVLQNRVHRTEEVKHVDFYAFSYYYDL 388

Qy        336 AVDTDMIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395
              |    :|| |||| | | |||  |: ||  ||     ||| ||||:|:: ||:: |||
Db        389 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTLETQPQSSPFSCMDLTYVSLLLQE-FGFPR 447

Qy        396 STVLQLTKKVNNIETGWALGATFHLLQSL 424
              | ||:||:|::|:|| ||||| || : ||
Db        448 SKVLKLTRKIDNVETSWALGAIFHYIDSL 476


RESULT 13
ABB06124
ID   ABB06124 standard; Protein; 463 AA.
XX
AC   ABB06124;
XX
DT   10-MAY-2002  (first entry)
XX
DE   Human NS protein sequence SEQ ID NO:216.
XX
KW   Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW   antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW   vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW   anorectic; muscular; antiinfertility; cardiovascular; anticoagulant;
KW   antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KW   anticonvulsant; antidiabetic; tranquilliser; antidepressant; aeuroleptic;
KW   gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
KW   contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW   endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW   rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW   inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW   infertility; cardiovascular disease; coagulation disease; hypertension;
KW   ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KW   diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW   gastric ulcer; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   WO200206315-A2.
XX
PD   24-JAN-2002.
XX
PF   17-JUL-2001; 2001WO-IL00653.
XX
PR   18-JUL-2000; 2000IL-0137345.
PR   15-DEC-2000; 2000IL-0140354.
XX
PA   (COMP-) COMPUGEN LTD.
XX
PI   Mintz L,  Freilich S,  Bernstein J;
XX
DR   WPI; 2002-155037/20.
DR   N-PSDB; ABL39778.
XX
PT   One hundred and twenty eight novel nucleic acid sequences, useful for
PT   treating and diagnosing e.g. cancer, asthma and Alzheimer's -
XX
PS   Claim 6; Page 251-253; 290pp; English.
XX
CC   ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC   encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC   (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC   antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC   vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC   anorectic, muscular, anti-HIV, antiinfertility, cardiovascular,
CC   anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC   immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
CC   antidepressant, gastrointestinal, aeuroleptic, cerebroprotective,
CC   nootropic and contraceptive activities. The NS can be used in vaccines,
CC   gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC   antibodies from the present invention can be used for treating and
CC   diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC   diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC   cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC   glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC   disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC   disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC   depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC   Alzheimer's disease and as a contraceptive.
XX
SQ   Sequence   463 AA;

  Query Match             44.1%;  Score 993;  DB 23;  Length 463;
  Best Local Similarity   52.2%;  Pred. No. 5.2e-87;
  Matches  203;  Conservative   55;  Mismatches  125;  Indels    6;  Gaps    4;

Qy         40 PINVSA---STLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQ 96
              |:  :|      ||||||||||||:||: |  : | : | |  | | :|||||||: |
Db         69 PLGTAADGHEVFYGIMFDAGSTGTRVHVFQFT-RPPRETPTLTHETFKALKPGLSAYADD 127

Qy         97 PKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSP 156
              :: |: :: ||:|||  ||   || ||:|||||||||||||  ||: || :|||:|: ||
Db        128 VEKSAQGIRELLDVAKQDIPFDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASP 187

Qy        157 FLVPKGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKT 216
```

```
Db        188 FLVGDDCVSIMNGTDEGVSAWITINFLTGSLKTPGGSSVGMLDLGGGSTQIAFLPRVEGT 247

Qy        217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPR 275

Db        248 LQASPPGYLTALRMFNRTYKLYSYSYLGLGLMSARLAILGGVEGQPAKDGKELVSPCLSP 307

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDR 335

Db        308 SFKGEWEHAEVTYRVSGQKAAASLHELCAARVSEVLQNRVHRTEEVKHVDFYAFSYYYDL 367

Qy        336 AVDTDMIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395

Db        368 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTLETQPQSSPFSCMDLTYVSLLLQE-FGFPR 426

Qy        396 STVLQLTKKVNNIETGWALGATFHLLQSL 424

Db        427 SKVLKLTRKIDNVETSWALGAIFHYIDSL 455


RESULT 14
ABJ04657
ID   ABJ04657 standard; Protein; 467 AA.
XX
AC   ABJ04657;
XX
DT   11-OCT-2002  (first entry)
XX
DE   Protein of NOVX 15a SEQ ID No 36.
XX
KW   Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antilipaemic;
KW   neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
KW   tranquiliser; neuroleptic; antidiabetic; antiulcer; antiinflammatory;
KW   anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;
KW   metabolic disorder; obesity; infectious disease; Alzheimer's disease;
KW   anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;
KW   immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
KW   metabolic syndrome X; wasting disorder; cancer; neurological disorder;
KW   epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;
KW   vesicular transport; cystic fibrosis; gastrointestinal disorder;
KW   diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
KW   multiple sclerosis; rheumatoid arthritis; transgenic animal;
KW   gene therapy.
XX
OS   Unidentified.
XX
PN   WO200246409-A2.
XX
PD   13-JUN-2002.
XX
PF   06-DEC-2001; 2001WO-US46586.
XX
PR   06-DEC-2000; 2000US-251660P.
PR   12-DEC-2000; 2000US-255029P.
PR   08-JAN-2001; 2001US-260326P.
PR   24-JAN-2001; 2001US-263800P.
PR   20-FEB-2001; 2001US-269942P.
PR   24-APR-2001; 2001US-286183P.
PR   20-AUG-2001; 2001US-313627P.
PR   12-SEP-2001; 2001US-318712P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Guo X,  Li L,  Patturajan M,  Shimkets RA,  Casman SJ,  Malyankar UM;
PI   Tchernev VT,  Vernet CAM,  Spytek KA,  Shenoy SG,  Alsobrook JP;
PI   Edinger S,  Peyman JA,  Stone DJ,  Ellerman K,  Gangolli EA;
PI   Boldog FL,  Colman SD,  Eisen AJ,  Liu X,  Padigaru M,  Spaderna SK;
PI   Zerhusen BD;
XX
DR   WPI; 2002-547774/58.
DR   N-PSDB; ABT05470.
XX
PT   Novel isolated polypeptide, designated NOVX, useful for treating or
PT   preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
PT   metabolic, neurodegenerative, immune and hematopoietic disorders  -
XX
PS   Claim 1; Page 140; 421pp; English.
XX
CC   The invention relates to an isolated polypeptide, designated NOVX,
CC   comprising a sequence fully defined in the specification. The isolated
CC   protein, its encoding polynucleotide or an antibody created from the
CC   protein is useful in the manufacture of a medicament for treating a
CC   syndrome associated with a human disease, preferably a NOVX-associated
CC   disorder, or for treating or preventing a NOVX-associated disorder in a
CC   subject, preferably human. The isolated protein, its encoding
CC   polynucleotide or an antibody created from the protein are also useful
CC   for treating or preventing metabolic disorders, diabetes, obesity,
CC   infectious disease, anorexia, neurodegenerative disorder, Alzheimer's
CC   disease, Parkinson's disorder, immune disorders, haematopoietic
CC   disorders, and various dyslipidaemias, metabolic disturbances associated
CC   with obesity, the metabolic syndrome X, wasting disorders associated with
CC   chronic diseases, and cancer. The isolated protein, its encoding
CC   polynucleotide or an antibody created from the protein are useful for
CC   treating or preventing neurological disorders such as epilepsy, stroke,
CC   mental disorders including mood, anxiety, schizophrenic disorders,
CC   disorders of vesicular transport such as cystic fibrosis, diabetes
CC   mellitus, goiter, gastrointestinal disorders including ulcerative
CC   colitis, other conditions associated with abnormal vesicle trafficking
CC   including AIDS, allergic reactions, multiple sclerosis and rheumatoid
CC   arthritis. A cell comprising the vector of the invention is useful for
CC   producing non-human transgenic animals. The polynucleotide of the
CC   invention can be used to treat disorders by gene therapy. This sequence
CC   represents one of the isolated NOVX proteins of the invention.
XX
SQ   Sequence   467 AA;

  Query Match             44.1%;  Score 993;  DB 23;  Length 467;
  Best Local Similarity   52.2%;  Pred. No. 5.2e-87;
  Matches  203;  Conservative   55;  Mismatches  125;  Indels    6;  Gaps    4;

Qy         40 PINVSA---STLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQ 96

Db         73 PLGTAADGHEVFYGIMFDAGSTGTRVHVFQFT-RPPRETPTLTHETFKALKPGLSAYADD 131

Qy         97 PKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSP 156

Db        132 VEKSAQGIRELLDVAKQDIPFDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASP 191

Qy        157 FLVPKGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKT 216

Db        192 FLVGDDCVSIMNGTDEGVSAWITINFLTGSLKTPGGSSVGMLDLGGGSTQIAFLPRVEGT 251

Qy        217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPR 275

Db        252 LQASPPGYLTALRMFNRTYKLYSYSYLGLGLMSARLAILGGVEGQPAKDGKELVSPCLSP 311

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDR 335

Db        312 SFKGEWEHAEVTYRVSGQKAAASLHELCAARVSEVLQNRVHRTEEVKHVDFYAFSYYYDL 371

Qy        336 AVDTDMIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395

Db        372 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTLETQPQSSPFSCMDLTYVSLLLQE-FGFPR 430

Qy        396 STVLQLTKKVNNIETGWALGATFHLLQSL 424

Db        431 SKVLKLTRKIDNVETSWALGAIFHYIDSL 459


RESULT 15
AAM93929
ID   AAM93929 standard; Protein; 456 AA.
XX
AC   AAM93929;
XX
DT   06-NOV-2001  (first entry)
```

XX
DE Human polypeptide, SEQ ID NO: 4100.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
DR N-PSDB; AAK94892.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 4100; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA ...
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 456 AA;

Query Match 43.9%; Score 989; DB 22; Length 456;
Best Local Similarity 51.9%; Pred. No. 1.2e-86;
Matches 202; Conservative 55; Mismatches 126; Indels 6; Gaps 4;

```
Qy         40 PINVSA---STLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQ 96
              |:  :|      |||||||||||||:||: |  : | : | |  | | ::||||||: |
Db         62 PLGTAADGHEVFYGIMFDAGSTGTRVHVFQFT-RPPRETPTLTHETFKALKPGLSAYADD 120

Qy         97 PKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSP 156
               :: |: :: ||:|||  ||   || ||:||||| ||||||  ||: || :|||:|: ||
Db        121 VEKSAQGIRELLDVAKQDIPFDFWKATPLVLKATVGLRLLPGEKAQKLLQKVKEVFKASP 180

Qy        157 FLVPKGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKT 216
              |||    |||| | |||: ||:|:||||| |      :|| ||||| |||| |||: | |
Db        181 FLVGDDCVSIMNGTDEGVSAWITINFLTGSLKTPGGSSVGMLDLGGGSTQIAFLPRVEGT 240

Qy        217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPR 275
              |: :| ||||:  ||| ||||::|||| || :|||| || :| :   ||    | ||
Db        241 LQASPPGYLTALRMFNRTYKLYSYSYLGLGLMSARLAILGGVEGQPAKDGKELVSPCLSP 300

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDR 335
               : ||   | |:  | :      | | | |  |:: ::|: |||:   |||||||||
Db        301 SFKGEWEHAEVTYRVSGQKAAASLHELCAARVSEVLQNRVHRTEEVKHVDFYAFSYYYDL 360

Qy        336 AVDTDMIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395
              |   :|| |||| | | |||  |: ||  ||     ||| ||||:|:: ||:: |||
Db        361 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTLETQPQSSPFSCMDLTYVSLLLQE-FGFPR 419
```

```
Qy        396 STVLQLTKKVNNIETGWALGATFHLLQSL 424
              | ||:||:|::|:|| ||||| || : ||
Db        420 SKVLKLTRKIDNVETSWALGAIFHYIDSL 448
```

Search completed: January 29, 2004, 10:03:37
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 10:02:15 ; Search time 15 Seconds
(without alignments)
1207.270 Million cell updates/sec

Title: US-10-091-085-7
Perfect score: 2252
Sequence: 1 MATSWGTVFFMLVVSCVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2252	100.0	428	4	US-09-608-285A-7	Sequence 7, Appli
2	2252	100.0	428	4	US-09-350-836B-7	Sequence 7, Appli
3	2252	100.0	428	4	US-09-370-265-7	Sequence 7, Appli
4	2252	100.0	428	4	US-09-557-800C-7	Sequence 7, Appli
5	2235	99.2	428	4	US-09-608-285A-3	Sequence 3, Appli
6	2235	99.2	428	4	US-09-608-285A-5	Sequence 5, Appli
7	2235	99.2	428	4	US-09-240-639-6	Sequence 6, Appli
8	2235	99.2	428	4	US-09-240-639-9	Sequence 9, Appli
9	2235	99.2	428	4	US-09-350-836B-3	Sequence 3, Appli
10	2235	99.2	428	4	US-09-350-836B-5	Sequence 5, Appli
11	2235	99.2	428	4	US-09-370-265-3	Sequence 3, Appli
12	2235	99.2	428	4	US-09-370-265-5	Sequence 5, Appli
13	2235	99.2	428	4	US-09-557-800C-3	Sequence 3, Appli
14	2235	99.2	428	4	US-09-557-800C-5	Sequence 5, Appli
15	2089	92.8	405	4	US-09-608-285A-25	Sequence 25, Appl
16	2089	92.8	405	4	US-09-370-265-25	Sequence 25, Appl
17	2089	92.8	405	4	US-09-557-800C-25	Sequence 25, Appl
18	1822.5	80.9	465	4	US-09-557-800C-56	Sequence 56, Appl
19	1817.5	80.7	465	4	US-09-240-639-8	Sequence 8, Appli
20	996	44.2	456	4	US-09-240-639-2	Sequence 2, Appli
21	996	44.2	484	4	US-09-608-285A-27	Sequence 27, Appl
22	996	44.2	484	4	US-09-370-265-27	Sequence 27, Appl
23	996	44.2	484	4	US-09-557-800C-27	Sequence 27, Appl
24	816.5	36.3	471	4	US-09-608-285A-60	Sequence 60, Appl
25	507.5	22.5	467	4	US-09-129-112-19	Sequence 19, Appl
26	502.5	22.3	459	4	US-09-129-112-9	Sequence 9, Appli
27	493	21.9	462	4	US-09-129-112-2	Sequence 2, Appli
28	486.5	21.6	462	4	US-09-129-112-15	Sequence 15, Appl
29	484	21.5	455	4	US-09-240-639-10	Sequence 10, Appl
30	472.5	21.0	473	4	US-09-240-639-12	Sequence 12, Appl
31	441.5	19.6	454	4	US-09-240-639-11	Sequence 11, Appl
32	366	16.3	502	4	US-09-557-800C-55	Sequence 55, Appl
33	366	16.3	510	3	US-08-930-921-1	Sequence 1, Appli
34	357.5	15.9	529	4	US-09-240-639-4	Sequence 4, Appli
35	229	10.2	148	4	US-09-240-639-17	Sequence 17, Appl
36	201	8.9	153	4	US-09-240-639-13	Sequence 13, Appl
37	191	8.5	153	4	US-09-240-639-15	Sequence 15, Appl
38	189.5	8.4	150	4	US-09-240-639-16	Sequence 16, Appl
39	189	8.4	154	4	US-09-240-639-14	Sequence 14, Appl
40	106.5	4.7	339	4	US-09-252-991A-20854	Sequence 20854, A
41	104.5	4.6	553	1	US-08-565-386-12	Sequence 12, Appl
42	103.5	4.6	535	3	US-09-137-077-2	Sequence 2, Appli
43	95.5	4.2	510	4	US-09-252-991A-17138	Sequence 17138, A
44	94	4.2	506	4	US-09-134-001C-4383	Sequence 4383, Ap
45	91.5	4.1	334	4	US-09-252-991A-22395	Sequence 22395, A

ALIGNMENTS

RESULT 1
US-09-608-285A-7
; Sequence 7, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-285A-7

Query Match 100.0%; Score 2252; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 4.4e-247;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
```

```
Db        61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy       121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Db       121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy       181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Db       181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy       241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Db       241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy       301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Db       301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy       361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Db       361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy       421 LQSLGISH 428

Db       421 LQSLGISH 428
```

RESULT 2
US-09-350-836B-7
; Sequence 7, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-836B-7

Query Match 100.0%; Score 2252; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 4.4e-247;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Db         1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy        61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Db        61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy       121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Db       121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy       181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Db       181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy       241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Db       241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy       301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Db       301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy       361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Db       361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy       421 LQSLGISH 428

Db       421 LQSLGISH 428
```

RESULT 3
US-09-370-265-7
; Sequence 7, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28111/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: PCT/US99/16180
; EARLIER FILING DATE: 1999-07-16
; EARLIER APPLICATION NUMBER: 09/350,836
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/273,447
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/244,444
; EARLIER FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 09/122,449
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: 09/118,205
; EARLIER FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-370-265-7

Query Match 100.0%; Score 2252; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 4.4e-247;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Db         1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy        61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Db        61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy       121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Db       121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy       181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Db       181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy       241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
                ||||||||
Db        421 LQSLGISH 428


RESULT 4
US-09-557-800C-7
; Sequence 7, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  APPLICANT:  Yeung, George
;  TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 28110/36457
;  CURRENT APPLICATION NUMBER: US/09/557,800C
;  CURRENT FILING DATE:  2000-04-25
;  PRIOR APPLICATION NUMBER: 09/481,238
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 1999-08-09
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/122449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/244444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-557-800C-7

  Query Match           100.0%;  Score 2252;  DB 4;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 4.4e-247;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
                ||||||||
Db        421 LQSLGISH 428


RESULT 5
US-09-608-285A-3
; Sequence 3, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
;  APPLICANT:  Ford, John
;  APPLICANT:  Mulero, Julio
;  APPLICANT:  Yeung, George
;  TITLE OF INVENTION:  METHODS AND MATERIALS RELATING TO CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE:  28110/36570
;  CURRENT APPLICATION NUMBER: US/09/608,285A
;  CURRENT FILING DATE:  2000-06-30
;  PRIOR APPLICATION NUMBER:  09/583,231
;  PRIOR FILING DATE:  2000-05-26
;  PRIOR APPLICATION NUMBER:  09/557,800
;  PRIOR FILING DATE:  2000-04-25
;  PRIOR APPLICATION NUMBER:  09/481,238
;  PRIOR FILING DATE:  2000-01-11
;  PRIOR APPLICATION NUMBER:  09/370,265
;  PRIOR FILING DATE:  1999-08-09
;  PRIOR APPLICATION NUMBER:  PCT/US99/16180
;  PRIOR FILING DATE:  1999-07-16
;  PRIOR APPLICATION NUMBER:  09/350,836
;  PRIOR FILING DATE:  1999-07-09
;  PRIOR APPLICATION NUMBER:  09/273,447
;  PRIOR FILING DATE:  1999-03-19
;  PRIOR APPLICATION NUMBER:  09/244,444
;  PRIOR FILING DATE:  1999-02-04
;  PRIOR APPLICATION NUMBER:  09/122,449
;  PRIOR FILING DATE:  1998-07-24
;  PRIOR APPLICATION NUMBER:  09/118,205
;  PRIOR FILING DATE:  1998-07-16
;  NUMBER OF SEQ ID NOS:  60
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-608-285A-3

  Query Match           99.2%;  Score 2235;  DB 4;  Length 428;
  Best Local Similarity  99.3%;  Pred. No. 3.7e-245;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
```

```
Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 6
US-09-608-285A-5
; Sequence 5, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-285A-5

Query Match 99.2%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 3.7e-245;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 7
US-09-240-639-6
; Sequence 6, Application US/09240639
; Patent No. 6350447
; GENERAL INFORMATION:
; APPLICANT: Chadwick, Brian Paul
; APPLICANT: Frischauf, Anna-Maria
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES AND NUCLEIC ACIDS
; FILE REFERENCE: 9598-066
; CURRENT APPLICATION NUMBER: US/09/240,639
; CURRENT FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-240-639-6

Query Match 99.2%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 3.7e-245;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 8
US-09-240-639-9
; Sequence 9, Application US/09240639
; Patent No. 6350447
; GENERAL INFORMATION:
;  APPLICANT: Chadwick, Brian Paul
;  APPLICANT: Frischauf, Anna-Maria
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
;  TITLE OF INVENTION: POLYPEPTIDES AND NUCLEIC ACIDS
;  FILE REFERENCE: 9598-066
;  CURRENT APPLICATION NUMBER: US/09/240,639
;  CURRENT FILING DATE: 1998-01-29
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-240-639-9

  Query Match           99.2%;  Score 2235;  DB 4;  Length 428;
  Best Local Similarity  99.3%;  Pred. No. 3.7e-245;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 9
US-09-350-836B-3
; Sequence 3, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
;  APPLICANT:  Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION:  METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE:  28110/35761
;  CURRENT APPLICATION NUMBER: US/09/350,836B
;  CURRENT FILING DATE:  1999-07-09
;  PRIOR APPLICATION NUMBER:  09/273,447
;  PRIOR FILING DATE:  1999-03-19
;  PRIOR APPLICATION NUMBER:  09/118,205
;  PRIOR FILING DATE:  1998-07-16
;  PRIOR APPLICATION NUMBER:  09/122,449
;  PRIOR FILING DATE:  1998-07-24
;  PRIOR APPLICATION NUMBER:  09/244,444
;  PRIOR FILING DATE:  1999-02-04
;  NUMBER OF SEQ ID NOS:  23
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-350-836B-3

  Query Match           99.2%;  Score 2235;  DB 4;  Length 428;
  Best Local Similarity  99.3%;  Pred. No. 3.7e-245;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 10
US-09-350-836B-5
; Sequence 5, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
;  APPLICANT:  Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION:  METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE:  28110/35761
;  CURRENT APPLICATION NUMBER: US/09/350,836B
;  CURRENT FILING DATE:  1999-07-09
;  PRIOR APPLICATION NUMBER:  09/273,447
```

```
;  PRIOR FILING DATE:  1999-03-19
;  PRIOR APPLICATION NUMBER:  09/118,205
;  PRIOR FILING DATE:  1998-07-16
;  PRIOR APPLICATION NUMBER:  09/122,449
;  PRIOR FILING DATE:  1998-07-24
;  PRIOR APPLICATION NUMBER:  09/244,444
;  PRIOR FILING DATE:  1999-02-04
;  NUMBER OF SEQ ID NOS:  23
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-350-836B-5

  Query Match             99.2%;  Score 2235;  DB 4;  Length 428;
  Best Local Similarity   99.3%;  Pred. No. 3.7e-245;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 11
US-09-370-265-3
; Sequence 3, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT: Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 28111/35908
;  CURRENT APPLICATION NUMBER: US/09/370,265
;  CURRENT FILING DATE: 1999-08-09
;  EARLIER APPLICATION NUMBER: PCT/US99/16180
;  EARLIER FILING DATE: 1999-07-16
;  EARLIER APPLICATION NUMBER: 09/350,836
;  EARLIER FILING DATE: 1999-07-09
;  EARLIER APPLICATION NUMBER: 09/273,447
;  EARLIER FILING DATE: 1999-03-19
;  EARLIER APPLICATION NUMBER: 09/244,444
;  EARLIER FILING DATE: 1999-02-04
;  EARLIER APPLICATION NUMBER: 09/122,449
;  EARLIER FILING DATE: 1998-07-24
;  EARLIER APPLICATION NUMBER: 09/118,205
;  EARLIER FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-370-265-3

  Query Match             99.2%;  Score 2235;  DB 4;  Length 428;
  Best Local Similarity   99.3%;  Pred. No. 3.7e-245;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 12
US-09-370-265-5
; Sequence 5, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT: Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 28111/35908
;  CURRENT APPLICATION NUMBER: US/09/370,265
;  CURRENT FILING DATE: 1999-08-09
;  EARLIER APPLICATION NUMBER: PCT/US99/16180
;  EARLIER FILING DATE: 1999-07-16
;  EARLIER APPLICATION NUMBER: 09/350,836
;  EARLIER FILING DATE: 1999-07-09
;  EARLIER APPLICATION NUMBER: 09/273,447
;  EARLIER FILING DATE: 1999-03-19
;  EARLIER APPLICATION NUMBER: 09/244,444
;  EARLIER FILING DATE: 1999-02-04
;  EARLIER APPLICATION NUMBER: 09/122,449
;  EARLIER FILING DATE: 1998-07-24
;  EARLIER APPLICATION NUMBER: 09/118,205
;  EARLIER FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 37
```

```
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-370-265-5

  Query Match             99.2%;  Score 2235;  DB 4;  Length 428;
  Best Local Similarity   99.3%;  Pred. No. 3.7e-245;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 13
US-09-557-800C-3
; Sequence 3, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  APPLICANT:  Yeung, George
;  TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 28110/36457
;  CURRENT APPLICATION NUMBER: US/09/557,800C
;  CURRENT FILING DATE:  2000-04-25
;  PRIOR APPLICATION NUMBER: 09/481,238
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 1999-08-09
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/122449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/244444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-557-800C-3

  Query Match             99.2%;  Score 2235;  DB 4;  Length 428;
  Best Local Similarity   99.3%;  Pred. No. 3.7e-245;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 14
US-09-557-800C-5
; Sequence 5, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  APPLICANT:  Yeung, George
;  TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 28110/36457
;  CURRENT APPLICATION NUMBER: US/09/557,800C
;  CURRENT FILING DATE:  2000-04-25
;  PRIOR APPLICATION NUMBER: 09/481,238
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 1999-08-09
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/122449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/244444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/118,205
```

```
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-557-800C-5

  Query Match           99.2%;  Score 2235;  DB 4;  Length 428;
  Best Local Similarity  99.3%;  Pred. No. 3.7e-245;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 15
US-09-608-285A-25
; Sequence 25, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
;  APPLICANT:  Ford, John
;  APPLICANT:  Mulero, Julio
;  APPLICANT:  Yeung, George
;  TITLE OF INVENTION:  METHODS AND MATERIALS RELATING TO CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE:  28110/36570
;  CURRENT APPLICATION NUMBER: US/09/608,285A
;  CURRENT FILING DATE:  2000-06-30
;  PRIOR APPLICATION NUMBER:  09/583,231
;  PRIOR FILING DATE:  2000-05-26
;  PRIOR APPLICATION NUMBER:  09/557,800
;  PRIOR FILING DATE:  2000-04-25
;  PRIOR APPLICATION NUMBER:  09/481,238
;  PRIOR FILING DATE:  2000-01-11
;  PRIOR APPLICATION NUMBER:  09/370,265
;  PRIOR FILING DATE:  1999-08-09
;  PRIOR APPLICATION NUMBER:  PCT/US99/16180
;  PRIOR FILING DATE:  1999-07-16
;  PRIOR APPLICATION NUMBER:  09/350,836
;  PRIOR FILING DATE:  1999-07-09
;  PRIOR APPLICATION NUMBER:  09/273,447
;  PRIOR FILING DATE:  1999-03-19
;  PRIOR APPLICATION NUMBER:  09/244,444
;  PRIOR FILING DATE:  1999-02-04
;  PRIOR APPLICATION NUMBER:  09/122,449
;  PRIOR FILING DATE:  1998-07-24
;  PRIOR APPLICATION NUMBER:  09/118,205
;  PRIOR FILING DATE:  1998-07-16
;  NUMBER OF SEQ ID NOS:  60
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 25
;   LENGTH: 405
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-608-285A-25

  Query Match           92.8%;  Score 2089;  DB 4;  Length 405;
  Best Local Similarity  99.2%;  Pred. No. 1.4e-228;
  Matches  397;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
              |||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400


Search completed: January 29, 2004, 10:06:58
Job time : 16 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 10:02:15 ; Search time 15 Seconds
(without alignments)
1207.270 Million cell updates/sec

Title: US-10-091-085-3
Perfect score: 2250
Sequence: 1 MATSWGTVFFMLVVSCVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2250	100.0	428	4	US-09-608-285A-3	Sequence 3, Appli
2	2250	100.0	428	4	US-09-608-285A-5	Sequence 5, Appli
3	2250	100.0	428	4	US-09-240-639-6	Sequence 6, Appli
4	2250	100.0	428	4	US-09-240-639-9	Sequence 9, Appli
5	2250	100.0	428	4	US-09-350-836B-3	Sequence 3, Appli
6	2250	100.0	428	4	US-09-350-836B-5	Sequence 5, Appli
7	2250	100.0	428	4	US-09-370-265-3	Sequence 3, Appli
8	2250	100.0	428	4	US-09-370-265-5	Sequence 5, Appli
9	2250	100.0	428	4	US-09-557-800C-3	Sequence 3, Appli
10	2250	100.0	428	4	US-09-557-800C-5	Sequence 5, Appli
11	2235	99.3	428	4	US-09-608-285A-7	Sequence 7, Appli
12	2235	99.3	428	4	US-09-350-836B-7	Sequence 7, Appli
13	2235	99.3	428	4	US-09-370-265-7	Sequence 7, Appli
14	2235	99.3	428	4	US-09-557-800C-7	Sequence 7, Appli
15	2104	93.5	405	4	US-09-608-285A-25	Sequence 25, Appl
16	2104	93.5	405	4	US-09-370-265-25	Sequence 25, Appl
17	2104	93.5	405	4	US-09-557-800C-25	Sequence 25, Appl
18	1837.5	81.7	465	4	US-09-557-800C-56	Sequence 56, Appl
19	1832.5	81.4	465	4	US-09-240-639-8	Sequence 8, Appli
20	999	44.4	456	4	US-09-240-639-2	Sequence 2, Appli
21	999	44.4	484	4	US-09-608-285A-27	Sequence 27, Appl
22	999	44.4	484	4	US-09-370-265-27	Sequence 27, Appl
23	999	44.4	484	4	US-09-557-800C-27	Sequence 27, Appl
24	819.5	36.4	471	4	US-09-608-285A-60	Sequence 60, Appl
25	507.5	22.6	459	4	US-09-129-112-9	Sequence 9, Appli
26	503.5	22.4	467	4	US-09-129-112-19	Sequence 19, Appl
27	498	22.1	462	4	US-09-129-112-2	Sequence 2, Appli
28	491.5	21.8	462	4	US-09-129-112-15	Sequence 15, Appl
29	489	21.7	455	4	US-09-240-639-10	Sequence 10, Appl
30	467.5	20.8	473	4	US-09-240-639-12	Sequence 12, Appl
31	446.5	19.8	454	4	US-09-240-639-11	Sequence 11, Appl
32	351	15.6	502	4	US-09-557-800C-55	Sequence 55, Appl
33	351	15.6	510	3	US-08-930-921-1	Sequence 1, Appli
34	347.5	15.4	529	4	US-09-240-639-4	Sequence 4, Appli
35	234	10.4	148	4	US-09-240-639-17	Sequence 17, Appl
36	186	8.3	153	4	US-09-240-639-13	Sequence 13, Appl
37	180.5	8.0	150	4	US-09-240-639-16	Sequence 16, Appl
38	179	8.0	153	4	US-09-240-639-15	Sequence 15, Appl
39	174	7.7	154	4	US-09-240-639-14	Sequence 14, Appl
40	106.5	4.7	553	1	US-08-565-386-12	Sequence 12, Appl
41	105.5	4.7	339	4	US-09-252-991A-20854	Sequence 20854, A
42	98.5	4.4	535	3	US-09-137-077-2	Sequence 2, Appli
43	96	4.3	506	4	US-09-134-001C-4383	Sequence 4383, Ap
44	93	4.1	969	2	US-08-284-941-2	Sequence 2, Appli
45	93	4.1	969	2	US-08-447-642-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-608-285A-3
; Sequence 3, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-285A-3

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.9e-248;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
                 ||||||||
Db        421 LQSLGISH 428
```

RESULT 2
US-09-608-285A-5
; Sequence 5, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-285A-5

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.9e-248;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 3
US-09-240-639-6
; Sequence 6, Application US/09240639
; Patent No. 6350447
; GENERAL INFORMATION:
; APPLICANT: Chadwick, Brian Paul
; APPLICANT: Frischauf, Anna-Maria
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES AND NUCLEIC ACIDS
; FILE REFERENCE: 9598-066
; CURRENT APPLICATION NUMBER: US/09/240,639
; CURRENT FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-240-639-6

CD39L2 polypeptide

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.9e-248;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
```

```
Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENPTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

```
RESULT 4
US-09-240-639-9
; Sequence 9, Application US/09240639
; Patent No. 6350447
; GENERAL INFORMATION:
;  APPLICANT: Chadwick, Brian Paul
;  APPLICANT: Frischauf, Anna-Maria
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
;  TITLE OF INVENTION: POLYPEPTIDES AND NUCLEIC ACIDS
;  FILE REFERENCE: 9598-066
;  CURRENT APPLICATION NUMBER: US/09/240,639
;  CURRENT FILING DATE: 1998-01-29
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-240-639-9

  Query Match           100.0%;  Score 2250;  DB 4;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 3.9e-248;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

```
RESULT 5
US-09-350-836B-3
; Sequence 3, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
;  APPLICANT:  Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION:  METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE:  28110/35761
;  CURRENT APPLICATION NUMBER: US/09/350,836B
;  CURRENT FILING DATE:  1999-07-09
;  PRIOR APPLICATION NUMBER:  09/273,447
;  PRIOR FILING DATE:  1999-03-19
;  PRIOR APPLICATION NUMBER:  09/118,205
;  PRIOR FILING DATE:  1998-07-16
;  PRIOR APPLICATION NUMBER:  09/122,449
;  PRIOR FILING DATE:  1998-07-24
;  PRIOR APPLICATION NUMBER:  09/244,444
;  PRIOR FILING DATE:  1999-02-04
;  NUMBER OF SEQ ID NOS:  23
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-350-836B-3

  Query Match           100.0%;  Score 2250;  DB 4;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 3.9e-248;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

```
RESULT 6
US-09-350-836B-5
; Sequence 5, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
;  APPLICANT:  Ford, John
;  APPLICANT:  Mulero, Julio
```

```
;  TITLE OF INVENTION:  METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE:  28110/35761
;  CURRENT APPLICATION NUMBER: US/09/350,836B
;  CURRENT FILING DATE:  1999-07-09
;  PRIOR APPLICATION NUMBER:  09/273,447
;  PRIOR FILING DATE:  1999-03-19
;  PRIOR APPLICATION NUMBER:  09/118,205
;  PRIOR FILING DATE:  1998-07-16
;  PRIOR APPLICATION NUMBER:  09/122,449
;  PRIOR FILING DATE:  1998-07-24
;  PRIOR APPLICATION NUMBER:  09/244,444
;  PRIOR FILING DATE:  1999-02-04
;  NUMBER OF SEQ ID NOS:  23
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-350-836B-5

  Query Match           100.0%;  Score 2250;  DB 4;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 3.9e-248;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 7
US-09-370-265-3
; Sequence 3, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT: Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 28111/35908
;  CURRENT APPLICATION NUMBER: US/09/370,265
;  CURRENT FILING DATE: 1999-08-09
;  EARLIER APPLICATION NUMBER: PCT/US99/16180
;  EARLIER FILING DATE: 1999-07-16
;  EARLIER APPLICATION NUMBER: 09/350,836
;  EARLIER FILING DATE: 1999-07-09
;  EARLIER APPLICATION NUMBER: 09/273,447
;  EARLIER FILING DATE: 1999-03-19
;  EARLIER APPLICATION NUMBER: 09/244,444
;  EARLIER FILING DATE: 1999-02-04
;  EARLIER APPLICATION NUMBER: 09/122,449
;  EARLIER FILING DATE: 1998-07-24
;  EARLIER APPLICATION NUMBER: 09/118,205
;  EARLIER FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-370-265-3

  Query Match           100.0%;  Score 2250;  DB 4;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 3.9e-248;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 8
US-09-370-265-5
; Sequence 5, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT: Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION: POLYPEPTIDES
;  FILE REFERENCE: 28111/35908
;  CURRENT APPLICATION NUMBER: US/09/370,265
;  CURRENT FILING DATE: 1999-08-09
;  EARLIER APPLICATION NUMBER: PCT/US99/16180
;  EARLIER FILING DATE: 1999-07-16
;  EARLIER APPLICATION NUMBER: 09/350,836
;  EARLIER FILING DATE: 1999-07-09
;  EARLIER APPLICATION NUMBER: 09/273,447
;  EARLIER FILING DATE: 1999-03-19
;  EARLIER APPLICATION NUMBER: 09/244,444
```

; EARLIER FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 09/122,449
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: 09/118,205
; EARLIER FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-370-265-5

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.9e-248;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 9
US-09-557-800C-3
; Sequence 3, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457
; CURRENT APPLICATION NUMBER: US/09/557,800C
; CURRENT FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-800C-3

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.9e-248;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 10
US-09-557-800C-5
; Sequence 5, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457
; CURRENT APPLICATION NUMBER: US/09/557,800C
; CURRENT FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447

CD39L2 polynucleotide

```
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/122449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/244444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-557-800C-5

  Query Match            100.0%;  Score 2250;  DB 4;  Length 428;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-248;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 11
US-09-608-285A-7
; Sequence 7, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
;  APPLICANT:  Ford, John
;  APPLICANT:  Mulero, Julio
;  APPLICANT:  Yeung, George
;  TITLE OF INVENTION:  METHODS AND MATERIALS RELATING TO CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE:  28110/36570
;  CURRENT APPLICATION NUMBER: US/09/608,285A
;  CURRENT FILING DATE:  2000-06-30
;  PRIOR APPLICATION NUMBER:  09/583,231
;  PRIOR FILING DATE:  2000-05-26
;  PRIOR APPLICATION NUMBER:  09/557,800
;  PRIOR FILING DATE:  2000-04-25
;  PRIOR APPLICATION NUMBER:  09/481,238
;  PRIOR FILING DATE:  2000-01-11
;  PRIOR APPLICATION NUMBER:  09/370,265
;  PRIOR FILING DATE:  1999-08-09
;  PRIOR APPLICATION NUMBER:  PCT/US99/16180
;  PRIOR FILING DATE:  1999-07-16
;  PRIOR APPLICATION NUMBER:  09/350,836
;  PRIOR FILING DATE:  1999-07-09
;  PRIOR APPLICATION NUMBER:  09/273,447
;  PRIOR FILING DATE:  1999-03-19
;  PRIOR APPLICATION NUMBER:  09/244,444
;  PRIOR FILING DATE:  1999-02-04
;  PRIOR APPLICATION NUMBER:  09/122,449
;  PRIOR FILING DATE:  1998-07-24
;  PRIOR APPLICATION NUMBER:  09/118,205
;  PRIOR FILING DATE:  1998-07-16
;  NUMBER OF SEQ ID NOS:  60
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-608-285A-7

  Query Match            99.3%;  Score 2235;  DB 4;  Length 428;
  Best Local Similarity   99.3%;  Pred. No. 2e-246;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 12
US-09-350-836B-7
; Sequence 7, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
;  APPLICANT:  Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION:  METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE:  28110/35761
;  CURRENT APPLICATION NUMBER: US/09/350,836B
;  CURRENT FILING DATE:  1999-07-09
;  PRIOR APPLICATION NUMBER:  09/273,447
;  PRIOR FILING DATE:  1999-03-19
;  PRIOR APPLICATION NUMBER:  09/118,205
;  PRIOR FILING DATE:  1998-07-16
```

; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-836B-7

Query Match 99.3%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 2e-246;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||||| | |||| |||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 13
US-09-370-265-7
; Sequence 7, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28111/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: PCT/US99/16180
; EARLIER FILING DATE: 1999-07-16
; EARLIER APPLICATION NUMBER: 09/350,836
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/273,447
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/244,444
; EARLIER FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 09/122,449
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: 09/118,205
; EARLIER FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-370-265-7

Query Match 99.3%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 2e-246;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||||| | |||| |||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 14
US-09-557-800C-7
; Sequence 7, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457
; CURRENT APPLICATION NUMBER: US/09/557,800C
; CURRENT FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205

```
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;    LENGTH: 428
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-557-800C-7

  Query Match           99.3%;  Score 2235;  DB 4;  Length 428;
  Best Local Similarity   99.3%;  Pred. No. 2e-246;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 15
US-09-608-285A-25
; Sequence 25, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
;  APPLICANT:  Ford, John
;  APPLICANT:  Mulero, Julio
;  APPLICANT:  Yeung, George
;  TITLE OF INVENTION:  METHODS AND MATERIALS RELATING TO CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE:  28110/36570
;  CURRENT APPLICATION NUMBER: US/09/608,285A
;  CURRENT FILING DATE:  2000-06-30
;  PRIOR APPLICATION NUMBER:  09/583,231
;  PRIOR FILING DATE:  2000-05-26
;  PRIOR APPLICATION NUMBER:  09/557,800
;  PRIOR FILING DATE:  2000-04-25
;  PRIOR APPLICATION NUMBER:  09/481,238
;  PRIOR FILING DATE:  2000-01-11
;  PRIOR APPLICATION NUMBER:  09/370,265
;  PRIOR FILING DATE:  1999-08-09
;  PRIOR APPLICATION NUMBER:  PCT/US99/16180
;  PRIOR FILING DATE:  1999-07-16
;  PRIOR APPLICATION NUMBER:  09/350,836
;  PRIOR FILING DATE:  1999-07-09
;  PRIOR APPLICATION NUMBER:  09/273,447
;  PRIOR FILING DATE:  1999-03-19
;  PRIOR APPLICATION NUMBER:  09/244,444
;  PRIOR FILING DATE:  1999-02-04
;  PRIOR APPLICATION NUMBER:  09/122,449
;  PRIOR FILING DATE:  1998-07-24
;  PRIOR APPLICATION NUMBER:  09/118,205
;  PRIOR FILING DATE:  1998-07-16
;  NUMBER OF SEQ ID NOS:  60
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 25
;    LENGTH: 405
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-608-285A-25

  Query Match           93.5%;  Score 2104;  DB 4;  Length 405;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-231;
  Matches  400;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
              |||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400


Search completed: January 29, 2004, 10:06:57
Job time : 16 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:56:59 ; Search time 34 Seconds
(without alignments)
3248.427 Million cell updates/sec

Title: US-10-091-085-3
Perfect score: 2250
Sequence: 1 MATSWGTVFFMLVVSCVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2236	99.4	428	4	Q96RX0	Q96rx0 homo sapien
2	2104	93.5	407	4	Q8WUB3	Q8wub3 homo sapien
3	1990.5	88.5	427	11	Q8CD29	Q8cd29 mus musculu
4	1984.5	88.2	427	11	Q8BR23	Q8br23 mus musculu
5	996	44.3	483	4	Q8TAS7	Q8tas7 homo sapien
6	996	44.3	503	4	Q8N3H3	Q8n3h3 homo sapien
7	696.5	31.0	461	5	O76268	O76268 drosophila
8	696.5	31.0	464	5	Q9VQI8	Q9vqi8 drosophila
9	616.5	27.4	479	5	Q9XU84	Q9xu84 caenorhabdi
10	593	26.4	278	11	Q8CHZ3	Q8chz3 mus musculu
11	538.5	23.9	556	3	Q9UT35	Q9ut35 schizosacch
12	518.5	23.0	489	10	Q8H7L6	Q8h7l6 oryza sativ
13	515.5	22.9	599	3	Q8TGH6	Q8tgh6 candida alb
14	511	22.7	522	3	Q9HEM6	Q9hem6 kluyveromyc
15	507.5	22.6	455	10	Q9SPM6	Q9spm6 medicago sa
16	503.5	22.4	467	10	Q9SPM7	Q9spm7 dolichos bi
17	498	22.1	462	10	Q9XFC9	Q9xfc9 dolichos bi
18	491.5	21.8	456	10	Q9SPM8	Q9spm8 lotus japon
19	489	21.7	455	10	Q9FEA6	Q9fea6 pisum sativ
20	489	21.7	472	10	Q8L704	Q8l704 arabidopsis
21	488	21.7	463	10	Q9FVC3	Q9fvc3 glycine soj
22	488	21.7	472	10	Q9SPM5	Q9spm5 arabidopsis
23	486	21.6	455	10	Q8GTB1	Q8gtb1 pisum sativ
24	485	21.6	472	10	Q9M7B3	Q9m7b3 arabidopsis
25	484	21.5	471	10	Q9SQG2	Q9sqg2 arabidopsis
26	481	21.4	447	10	Q9SLV4	Q9slv4 pisum sativ
27	475	21.1	468	10	Q9FVC2	Q9fvc2 glycine soj
28	473	21.0	467	10	Q8RVT6	Q8rvt6 pisum sativ
29	472	21.0	407	10	Q9AVN8	Q9avn8 pisum sativ
30	472	21.0	455	10	Q8RVU0	Q8rvu0 pisum sativ
31	465	20.7	447	10	Q9FUI1	Q9fui1 pisum sativ
32	464.5	20.6	558	3	Q8TGG8	Q8tgg8 aspergillus
33	462	20.5	466	10	Q9AU15	Q9au15 medicago tr
34	449	20.0	455	10	Q8RVT9	Q8rvt9 pisum sativ
35	443	19.7	473	10	Q8RVT8	Q8rvt8 pisum sativ
36	436.5	19.4	556	3	Q9C2M0	Q9c2m0 neurospora
37	435.5	19.4	271	5	Q8IPZ6	Q8ipz6 drosophila
38	432.5	19.2	555	10	Q94EZ2	Q94ez2 arabidopsis
39	429.5	19.1	555	10	O80612	O80612 arabidopsis
40	424.5	18.9	224	11	Q8CDB6	Q8cdb6 mus musculu
41	397	17.6	483	10	Q9XI62	Q9xi62 arabidopsis
42	392	17.4	493	13	Q90X66	Q90x66 gallus gall
43	385	17.1	690	5	Q9BHV5	Q9bhv5 leishmania
44	370	16.4	539	11	Q8CDV7	Q8cdv7 mus musculu
45	367	16.3	510	11	Q921Q6	Q921q6 mus musculu

ALIGNMENTS

RESULT 1
Q96RX0
ID Q96RX0 PRELIMINARY; PRT; 428 AA.
AC Q96RX0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Pcph proto-oncogene protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20173601; PubMed=10708485;
RA Recio J.A., Zambrano N., Pena, Ld, Reig J.A., Rhoads A., Rouzaut A.,
RA Notario V.;
RT "The human PCPH proto-oncogene: cDNA identification, primary
RT structure, chromosomal mapping, and expression in normal and tumor
RT cells.";
RL Mol. Carcinog. 27:229-236(2000).
DR EMBL; AF136572; AAK82950.1; -.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
SQ SEQUENCE 428 AA; 47431 MW; F2C4F7DE650A44F6 CRC64;

Query Match 99.4%; Score 2236; DB 4; Length 428;
Best Local Similarity 99.5%; Pred. No. 3.4e-185;
Matches 426; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKAAGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
```

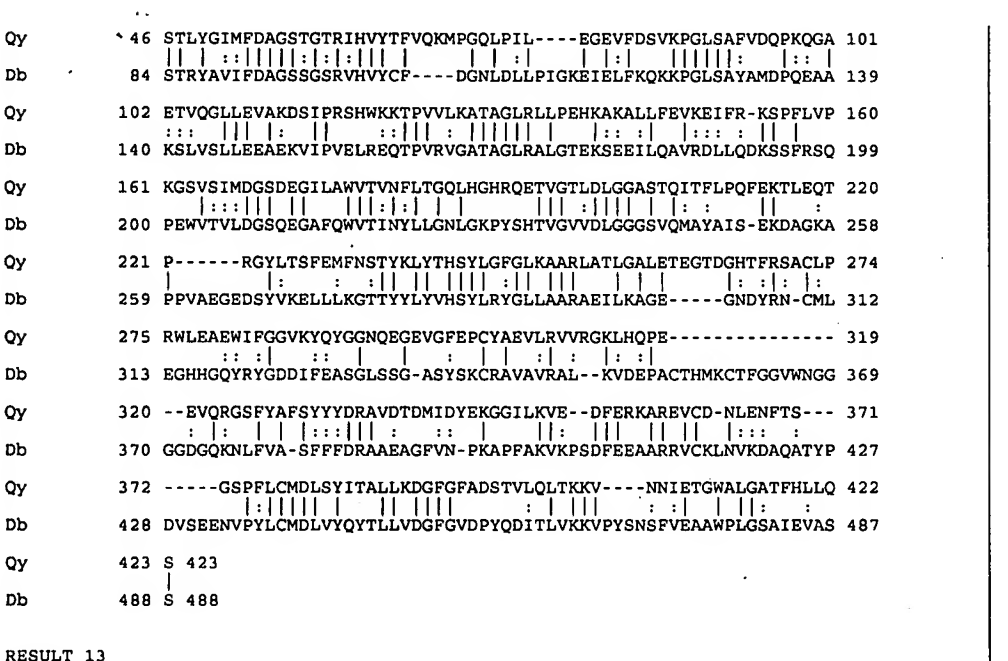
```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
                ||||||||
Db        421 LQSLGISH 428


RESULT 2
Q8WUB3
ID   Q8WUB3      PRELIMINARY;      PRT;   407 AA.
AC   Q8WUB3;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to ectonucleoside triphosphate diphosphohydrolase 5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Colon;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC020966; AAH20966.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
KW   Hydrolase.
SQ   SEQUENCE   407 AA;  45336 MW;  D92A5F7DC9EC9E5B CRC64;

  Query Match             93.5%;  Score 2104;  DB 4;  Length 407;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-174;
  Matches  400;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
                |||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400


RESULT 3
Q8CD29
ID   Q8CD29      PRELIMINARY;      PRT;   427 AA.
AC   Q8CD29;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK031581; BAC27461.1; -.
SQ   SEQUENCE   427 AA;  47101 MW;  6E3773C842B58477 CRC64;

  Query Match             88.5%;  Score 1990.5;  DB 11;  Length 427;
  Best Local Similarity   88.1%;  Pred. No. 6.5e-164;
  Matches  376;  Conservative   25;  Mismatches   25;  Indels    1;  Gaps    1;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
                |||||| | |||:::|| | | :| |||||||:||||||||||||| | ||||||||||||
Db          1 MATSWGAV-FMLIIACVGSTVFYREQQTWFEGVFLSSMCPINVSAGTFYGIMFDAGSTGT 59

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
                |||||||||| |||| ||||:||||||||||||||||||||| ||||||||||||||
Db         60 RIHVYTFVQKTAGQLPFLEGEIFDSVKPGLSAFVDQPKQGAETVQELLEVAKDSIPRSHW 119

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
                ::|||||||||||||||| ||:||| ||:|||: |||||| ||||||||| |||||||||
Db        120 ERTPVVLKATAGLRLLPEQKAQALLLEVEEIFKNSPFLVPDGSVSIMDGSYEGILAWVTV 179

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
                |||||||||  |||||||||||||||||||||||||:|||||||||||||||:|||||
Db        180 NFLTGQLHGHRGQETVGTLDLGGASTQITFLPQFEKTLEKTPRGYLTSFEMFNSTFKLYTH 239

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
                |||||||||||||||||||| :|||||||||||||||||||||||||||||||||||:||
Db        240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGF 299

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
                |||||||||||:|||||||:  :|||||||||| || :|||||:||||||||||||
Db        300 EPCYAEVLRVVQGKLHQPEEVRGSAFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 359

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
                |||||| :|:|||||||||:|||||||||||||| |:|||||||||||||||||||||
Db        360 EVCDNLGSFSSGSPFLCMDLTYITALLKDGFGFADGTLLQLTKKVNNIETGWALGATFHL 419

Qy        421 LQSLGIS 427
                |||||||
Db        420 LQSLGIT 426


RESULT 4
Q8BR23
ID   Q8BR23      PRELIMINARY;      PRT;   427 AA.
AC   Q8BR23;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
```

```
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK045828; BAC32507.1; -.
SQ   SEQUENCE   427 AA;  47123 MW;  87BF2CC1CCC1FCB9 CRC64;

  Query Match             88.2%;  Score 1984.5;  DB 11;  Length 427;
  Best Local Similarity   87.8%;  Pred. No. 2.2e-163;
  Matches  375;  Conservative   24;  Mismatches   27;  Indels    1;  Gaps    1;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              |||||| | |||:::|| | | :| |||||||:||||||||||| | ||||||||||||
Db          1 MATSWGAV-FMLIIACVGSTVFYREQQTWFEGVFLSSMCPINVSAGTFYGIMFDAGSTGT 59

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              |||||||||| |||| ||||:|||||||||||||||||||||| ||||||||||||||
Db         60 RIHVYTFVQKTAGQLPFLEGEIFDSVKPGLSAFVDQPKQGAETVQELLEVAKDSIPRSHW 119

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ::||  |||||||||||| ||:||| ||:|||: |||||| |||||||| |||||||||
Db        120 ERTPXXLKATAGLRLLPEQKAQALLLEVEEIFKNSPFLVPDGSVSIMDGSYEGILAWVTV 179

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              |||||||| ||||||||||||||||||||||||||||||||||||||:|||||
Db        180 NFLTGQLHGRGQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTFKLYTH 239

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGP 300
              |||||||||||||||||| :|||||||||||||||||||||||||||||||||||:||
Db        240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGP 299

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||:|||||||||:  :|||||||||| || :|||||||:||||||||||
Db        300 EPCYAEVLRVVQGKLHQPEEVRGSAFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 359

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              |||||| :|:||||||||||:|||||||||||||| |:||||||||||||||||||||
Db        360 EVCDNLGSFSSGSPFLCMDLTYITALLKDGFGFADGTLLQLTKKVNNIETGWALGATFHL 419

Qy        421 LQSLGIS 427
              |||||||
Db        420 LQSLGIT 426


RESULT 5
Q8TAS7
ID   Q8TAS7      PRELIMINARY;      PRT;   483 AA.
AC   Q8TAS7;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 6 (Putative
DE   function).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC025980; AAH25980.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
KW   Hydrolase.
SQ   SEQUENCE   483 AA;  53119 MW;  A8E0E5035BCDCE8F CRC64;

  Query Match             44.3%;  Score 996;  DB 4;  Length 483;
  Best Local Similarity   52.2%;  Pred. No. 1.3e-77;
  Matches  203;  Conservative   57;  Mismatches  123;  Indels    6;  Gaps    4;

Qy         40 PINVSA---STLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQ 96
              |:  :|      |||||||||||||:||: |  : | : | |  | | ::||||||: |
Db         89 PLGTAADGHEVFYGIMFDAGSTGTRVHVFQFT-RPPRETPTLTHETFKALKPGLSAYADD 147

Qy         97 PKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSP 156
              :: |: :: ||:|||  ||   || ||:||||||||||||  ||: || :|||:|: ||
Db        148 VEKSAQGIRELLDVAKQDIPFDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASP 207

Qy        157 FLVPKGSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKT 216
              |||    ||||:|:|||: ||:|:||||| |       :|| ||||| |||| |||: | |
Db        208 FLVGDDCVSIMNGTDEGVSAWITINFLTGSLKTPGGSSVGMLDLGGGSTQIAFLPRVEGT 267

Qy        217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPR 275
              |: :| ||||:  ||| ||||||::|||| || :|||| || :| :   ||     | ||
Db        268 LQASPPGYLTALRMFNRTYKLYSYSYLGLGLMSARLAILGGVEGQPAKDGKELVSPCLSP 327

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDR 335
              : ||    | |:  | :      | | |  |:: ::|: |||:    |||||||||
Db        328 SFKGEWEHAEVTYRVSGQKAAASLHELCAARVSEVLQNRVHRTEEVKHVDFYAFSYYYDL 387

Qy        336 AVDTDMIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395
              |    :|| |||| | | |||  |: ||  ||     ||| ||||:|:: ||:: |||
Db        388 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTLETQPQSSPFSCMDLTYVSLLLQE-FGFPR 446

Qy        396 STVLQLTKKVNNIETGWALGATFHLLQSL 424
              | ||:||:|::|:|| ||||| || : ||
Db        447 SKVLKLTRKIDNVETSWALGAIFHYIDSL 475


RESULT 6
Q8N3H3
ID   Q8N3H3      PRELIMINARY;      PRT;   503 AA.
AC   Q8N3H3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   DKFZP761J1915.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Amygdala;
RA   Ansorge W., Wirkner U., Mewes H.W., Weil B., Wiemann S.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL834158; CAD38864.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 2.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   503 AA;  54763 MW;  163933F9139D8D9F CRC64;

  Query Match             44.3%;  Score 996;  DB 4;  Length 503;
  Best Local Similarity   52.2%;  Pred. No. 1.4e-77;
  Matches  203;  Conservative   57;  Mismatches  123;  Indels    6;  Gaps    4;

Qy         40 PINVSA---STLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQ 96
              |:  :|      |||||||||||||:||: |  : | : | |  | | ::||||||: |
Db        109 PLGTAADGHEVFYGIMFDAGSTGTRVHVFQFT-RPPRETPTLTHETFKALKPGLSAYADD 167
```

```
Qy          97 PKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSP 156
               :: |: :: ||:|||  ||   || ||:|||||||||||||  ||: || :|||:|: ||
Db         168 VEKSAQGIRELLDVAKQDIPFDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASP 227

Qy         157 FLVPKGSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKT 216
               |||    ||||:|:|||: ||:|:||||| |       :|| ||||| ||||| |||: | |
Db         228 FLVGDDCVSIMNGTDEGVSAWITINFLTGSLKTPGGSSVGMLDLGGGSTQIAFLPRVEGT 287

Qy         217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPR 275
               |: :| ||||:  ||| ||||::|||| || :|||| || :| :   ||    | || .
Db         288 LQASPPGYLTALRMFNRTYKLYSYSYLGLGLMSARLAILGGVEGQPAKDGKELVSPCLSP 347

Qy         276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDR 335
                 : ||    | |: | :       | | | |  |:: ::|: |||:   |||||||||
Db         348 SFKGEWEHAEVTYRVSGQKAAASLHELCAARVSEVLQNRVHRTEEVKHVDFYAFSYYYDL 407

Qy         336 AVDTDMIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395
               |    :|| |||| | | |||  |: ||  ||     ||| ||||:|:: ||:: |||
Db         408 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTLETQPQSSPFSCMDLTYVSLLLQE-FGFPR 466

Qy         396 STVLQLTKKVNNIETGWALGATFHLLQSL 424
               | ||:||:|::|:|| ||||| || : ||
Db         467 SKVLKLTRKIDNVETSWALGAIFHYIDSL 495


RESULT 7
O76268
ID   O76268      PRELIMINARY;      PRT;   461 AA.
AC   O76268;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NTPASE protein (LD11641P).
GN   NTPASE OR CG3059.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98341119; PubMed=9676430;
RA   Chadwick B.P., Frischauf A.-M.;
RT   "The CD39-like gene family: identification of three new human members
RT   (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of
RT   the gene family from Drosophila melanogaster.";
RL   Genomics 50:357-367(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003581; AAF51182.1; -.
DR   EMBL; AF041048; AAC39133.1; -.
DR   EMBL; AY061134; AAL28682.1; -.
DR   FlyBase; FBgn0024947; NTPase.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
SQ   SEQUENCE   461 AA;  50845 MW;  27D00321F91A9DD1 CRC64;

  Query Match             31.0%;  Score 696.5;  DB 5;  Length 461;
  Best Local Similarity   39.1%;  Pred. No. 1.1e-51;
  Matches  156;  Conservative   74;  Mismatches  122;  Indels   47;  Gaps   12;

Qy          49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLL 108
               |  : ||||||:|:  | | :       :|  |:|   |||||:| | | :|| ::: ||
Db          79 YAAIIDAGSTGSRVLAYKFNRSFIDNKLVLYEELFKERKPGLSSFADNPAEGAHSIKLLL 138

Qy         109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMD 168
               :.|:  ||: ||  ||:|||||||||||  ||: :|  |:::| || | |   :| |||
Db         139 DEARAFIPKEHWSSTPLVLKATAGLRLLPASKAENILNAVRDLFAKSEFSVDMDAVEIMD 198

Qy         169 GSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITF-------LPQFEKTLEQTP 221
               |:|||| :| ||||| |:|    |     ||||| |||:||       :| ::| : :
Db         199 GTDEGIFSWFTVNFLLGRLSKTNQ--AAALDLGGGSTQVTFSPTDPDQVPVYDKYMHEV- 255

Qy         222 RGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEG--TDGHTFRSACL-PRWLE 278
                  :|| :  |    ::|||||| || ||| |      | |   :      | |: |
Db         256 ---VTSSKKIN----VFTHSYLGLGLMAARHAVF----THGYKKEDTVLESVCVNPIIAN 304

Qy         279 AEWIFGGVKYQYGGNQEGE-------VGFEPCYAEVLRVVRGKL-----HQPEEVQRGSF 326
                 | :| |:|:  | : |:       | |: |    | :|: |:      :|  ::: :
Db         305 RTWTYGNVQYKVSGKENGKSSAEQPIVDFDAC----LELVKSKVMPLVKPKPFTLKQHAV 360

Qy         327 YAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAREVC--DNLENFTSGSPFLCMDLSYIT 384
               |||||::||::: ::|   ||   || : :||:|:|   | |      ||:| ||::|:
Db         361 AAFSYYFERAIESGLVDPLAGGETTVEAYRKKAQEICAIPNDE-----QPFMCFDLTFIS 415

Qy         385 ALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHLLQS 423
               ||::|||  |   ::| ||::  |  ||||  :::| |
Db         416 TLLREGFGLNDGKKIKLYKKIDGHEISWALGCAYNVLTS 454


RESULT 8
Q9VQI8
ID   Q9VQI8      PRELIMINARY;      PRT;   464 AA.
AC   Q9VQI8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
```

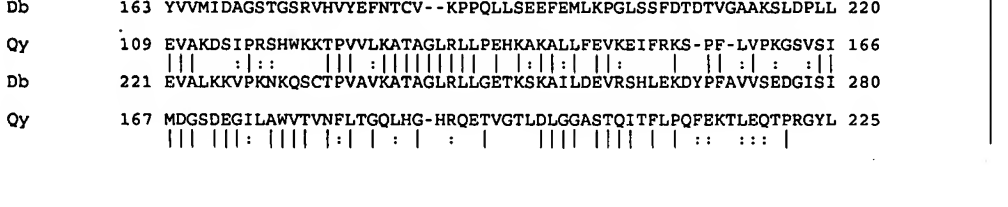
```
DE   NTPASE protein.
GN   NTPASE OR CG3059.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003581; AAF51181.1; -.
DR   FlyBase; FBgn0024947; NTPase.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
SQ   SEQUENCE   464 AA;  51119 MW;  71D057ABE5AE613D CRC64;

  Query Match           31.0%;  Score 696.5;  DB 5;  Length 464;
  Best Local Similarity  39.1%;  Pred. No. 1.2e-51;
  Matches  156;  Conservative   74;  Mismatches  122;  Indels   47;  Gaps   12;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLL 108
              |  : ||||||:|:  | | :       :|  |:|   |||||:| | | :|| ::: ||
Db         82 YAAIIDAGSTGSRVLAYKFNRSFIDNKLVLYEELFKERKPGLSSFADNPAEGAHSIKLLL 141

Qy        109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMD 168
              : |:  ||: ||  ||:|||||||||||||  ||: :|  |:::| || | |   :| |||
Db        142 DEARAFIPKEHWSSTPLVLKATAGLRLLPASKAENILNAVRDLFAKSEFSVDMDAVEIMD 201

Qy        169 GSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITF-------LPQFEKTLEQTP 221
              |:|||| :| ||||| |:|     |      ||||| |||:||       :| ::| : :
Db        202 GTDEGIFSWFTVNFLLGRLSKTNQ--AAALDLGGGSTQVTFSPTDPDQVPVYDKYMHEV- 258

Qy        222 RGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEG--TDGHTFRSACL-PRWLE 278
                 :|| :  |    ::|||||| || ||| |      | |   :     | |: |
Db        259 ---VTSSKKIN----VFTHSYLGLGLMAARHAVF----THGYKKEDTVLESVCVNPIIAN 307

Qy        279 AEWIFGGVKYQYGGNQEGE-------VGFEPCYAEVLRVVRGKL-----HQPEEVQRGSF 326
              | :| |:|:  | : |:       | |: |    | :|: |:      :|  ::: :
Db        308 RTWTYGNVQYKVSGKENGKSSAEQPIVDFDAC----LELVKSKVMPLVKPKPFTLKQHAV 363

Qy        327 YAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAREVC--DNLENFTSGSPFLCMDLSYIT 384
              |||||::||::: ::|   ||   || : :||:|:|   | |      ||:| ||::|:
Db        364 AAFSYYFERAIESGLVDPLAGGETTVEAYRKKAQEICAIPNDE-----QPFMCFDLTFIS 418

Qy        385 ALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHLLQS 423
              ||::|||  |   ::| ||::  |  ||||  :::| |
Db        419 TLLREGFGLNDGKKIKLYKKIDGHEISWALGCAYNVLTS 457


RESULT 9
Q9XU84
ID   Q9XU84      PRELIMINARY;      PRT;   479 AA.
AC   Q9XU84;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   K08H10.4 protein.
GN   K08H10.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gardner A.E.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z83113; CAB05544.1; -.
DR   WormPep; K08H10.4; CE18877.
DR   InterPro; IPR002034; AIPM/Hcit_synth.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS00815; AIPM_HOMOCIT_SYNTH_1; 1.
SQ   SEQUENCE   479 AA;  53844 MW;  7EDC02A9D54A48ED CRC64;

  Query Match           27.4%;  Score 616.5;  DB 5;  Length 479;
  Best Local Similarity  34.3%;  Pred. No. 1e-44;
  Matches  150;  Conservative   81;  Mismatches  167;  Indels   39;  Gaps   16;

Qy          9 FFMLVVS--CVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYT 66
              | :|::|   : |.|: : |  |  |  : :    |      : |: |||||||:|:|
Db          6 FSILLISFFSLLSVVTTKTQY-WCHGDGVLN----NQHTCRFFTIVIDAGSTGTRLHLYK 60

Qy         67 FVQK-------MPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSH 119
              |:          || :   :| |:|  ||||||:|    |   |:::: ||: |:  :|
Db         61 FIHDPAIASHGMPFK---VEKEIFQEVKPGLSSFAKSPSSAADSLEPLLQRARKEVPHFM 117

Qy        120 WKKTPVVLKATAGLRLLPEHKAKALLFEVKE-IFRKSPFLVPKGSVSIMDGSDEGILAWV 178
              |:|||: ||||||||||    | :|  |:| ||     :|::| |||||: :|
Db        118 WEKTPITLKATAGLRLLPGDMADDILESVEERIFNSGFFAAFPDAVNVMPGSDEGVYSWF 177

Qy        179 TVNFLTGQLH------GHR---QETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFE 229
              |:| |   |       ||:     :|   |||| |||:|: |  |     :|   ||   :
Db        178 TLNILLETLFTDEPTVGHKPAAHRSVAAFDLGGGSTQLTYWPNNEAVFSEHV-GYERDID 236

Qy        230 MFNSTYKLYTHSYLGFGLKAARLATLGALETEG--TDGHTFRSACLPRWLE-AEWIFGGV 286
              |    :|:|||:|| || ||||  |  |||:       |   ::|:|   :  || :  :
Db        237 FFGHHIRLFTHSFLGNGLIAARLNIL-QLETDNEIESTHQLITSCMPEGYQLTEWEY-AL 294

Qy        287 KYQYGGNQEGEVGFEPCYAEVLRVVR-GKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYE 345
              |: :  |      || ||      |  ::    |::    | |||::|||::: ::
```

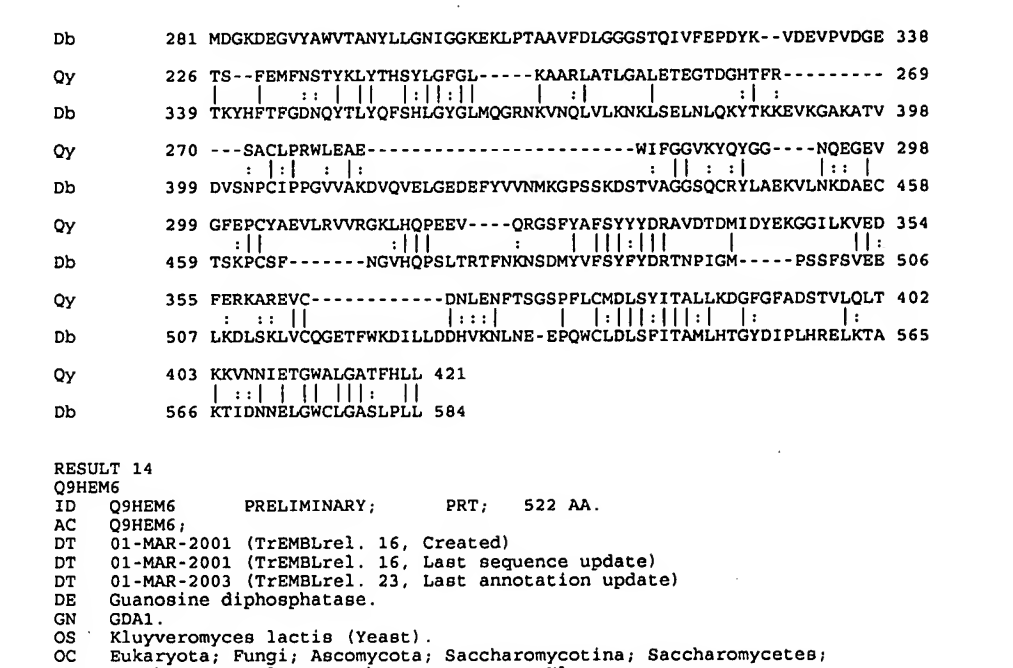
```
Db        295 KF-WNINGSSSHSFESCYGTTKNFVESSEIMHLRELKGSPVYLFSYFFDRALNSGLVKGN 353

Qy        346 KGGILKVEDFERKAREVC----DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQL 401
              :|| :::  |: |   |      ::: :   |: |:||:|| :||:||: | |:  | |
Db        354 EGGKIELRQFKEAAEIACRREKTEIDDGSHWMPWQCLDLTYIYSLLRDGYQFEDNQPLVL 413

Qy        402 TKKVNNIETGWALGATF 418
                ||:  :|  |  |  |
Db        414 AKKIKGMEVSWGQGLAF 430


RESULT 10
Q8CHZ3
ID   Q8CHZ3      PRELIMINARY;      PRT;   278 AA.
AC   Q8CHZ3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to ectonucleoside triphosphate diphosphohydrolase 6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA   Strausberg R.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC038126; AAH38126.1; -.
KW   Hydrolase.
SQ   SEQUENCE   278 AA;  30130 MW;  75A92DD1AC76297F CRC64;

  Query Match             26.4%;  Score 593;  DB 11;  Length 278;
  Best Local Similarity   58.0%;  Pred. No. 4.9e-43;
  Matches  119;  Conservative   31;  Mismatches   49;  Indels    6;  Gaps    2;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLL 108
              |||||||||||||||: |  : ||: | |  | | ::|||||: |  :: |: :| ||
Db         74 YGIMFDAGSTGTRIHVFQFA-RPPGETPTLTHETFKALKPGLSAYADDVEKSAQGIQELL 132

Qy        109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMD 168
               |||  ||   || ||:|||||||||||||  ||: || :|||:|: |||||    ||||:
Db        133 NVAKQHIPYDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASPFLVGDDCVSIMN 192

Qy        169 GSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSF 228
              |:|||: ||:|||||| |     :|| ||||| ||||||||: | ||: :| |:||:
Db        193 GTDEGVSAWITVNFLTGSLKTPGSSSVGMLDLGGGSTQITFLPRVEGTLQASPPGHLTAL 252

Qy        229 EMFNSTYKLYTHSYLGFGLKAARLA 253
              :||| |.||||:: ::      :|||
Db        253 QMFNRTYKLYSYRWV-----CSRLA 272


RESULT 11
Q9UT35
ID   Q9UT35      PRELIMINARY;      PRT;   556 AA.
AC   Q9UT35;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative guanosine-diphosphatase (Guanosine diphosphatase).
GN   SPAC824.08 OR GDP1.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972h-;
RA   Barrell B.G., Rajandream M.A., Quail M., Seegar K., Harris D.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Sanchez R., Franco A., Notario V., Gacto M., Cansado J.;
RT   "Characterization of a guanosine diphosphatase gene from
RT   Schizosaccharomyces pombe.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL121741; CAB57338.1; -.
DR   EMBL; AF465240; AAL69974.1; -.
DR   GeneDB_SPombe; SPAC824.08; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
SQ   SEQUENCE   556 AA;  61588 MW;  1D811E3D6A6BBB85 CRC64;

  Query Match             23.9%;  Score 538.5;  DB 3;  Length 556;
  Best Local Similarity   34.4%;  Pred. No. 7.7e-38;
  Matches  144;  Conservative   61;  Mismatches  157;  Indels   57;  Gaps   13;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLL 108
              | :| ||||||:|:||| |     |  | || | |  ::||||:|   |:  | ::  ||
Db        134 YVLMIDAGSTGSRVHVYQFNNCNPS--PKLEEEFFKMIEPGLSSFAGDPEGAAASLDPLL 191

Qy        109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKS-PFLVPKGSVSIM 167
              : | :::|  : : :|: :||||||||  | :|||:|  |::      || : |  |||:
Db        192 DYAMENVPEEYRRCSPIAVKATAGLRLTGESEAKAILKSVRQHLENDYPFPIVKDGVSIL 251

Qy        168 DGSDEGILAWVTVNFLTGQLHGH-RQETVGTLDLGGASTQITFLPQFEKTLEQTPRG--- 223
              :|| ||| ||:|:|:| | | |     ||  :|||||||||: | |:|    |    |
Db        252 EGSMEGIYAWITINYLLGTLGGKATHSTVAVMDLGGASTQLVFEPRFASDGESLVDGDHK 311

Qy        224 YLTSFEMFNSTYKLYTHSYLGFGLKAARL----------ATLGALETEGTDGHTFRSAC 272
              |: :      |:|| ||:|:||| ||             |   :||  | |  :    |
Db        312 YVLDYN--GEQYELYQHSHLGYGLKEARKLIHKFVLNNAEALKESLELLG-DSTSIIHPC 368

Qy        273 L----------PRWLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVR----GKLHQP 318
              |          :  :| |  |   :  |  : : | :    ||      :|||
Db        369 LHLNASLTHPDSKSEASEVVFVGPSLAHLSLQCRGIAEKALYKDKNCPVRPCSFNGVHQP 428

Qy        319 ---EEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAREVC------------ 363
                 |       | ||:|||     ||          :|| :  |  ||
Db        429 KFTETFTDSPIYLISYFYDR-----MISLGMPSTFTIEDMKYLANSVCSGPTYWQDAFSL 483

Qy        364 -DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHLL 421
               | |:   |  |:||:|: :||  |:   ::  |   ||::| | || |||:  :|
Db        484 TDALKELKE-EPEWCLDLNYMISLLSVGYEIPNNRQLHTAKKIDNKELGWCLGASLSML 541


RESULT 12
Q8H7L6
ID   Q8H7L6      PRELIMINARY;      PRT;   489 AA.
AC   Q8H7L6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative apyrase.
GN   OSJNBB0014I10.10.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saski C.,
RA   Currie J., Collura K.;
RT   "Rice Genomic Sequence.";
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC126222; AAN65004.1; -.
SQ   SEQUENCE   489 AA;  52799 MW;  0BF6BF154488D38E CRC64;

  Query Match             23.0%;  Score 518.5;  DB 10;  Length 489;
  Best Local Similarity   34.0%;  Pred. No. 3.4e-36;
  Matches  143;  Conservative   70;  Mismatches  149;  Indels   59;  Gaps   16;
```



RESULT 13
Q8TGH6
ID Q8TGH6 PRELIMINARY; PRT; 599 AA.
AC Q8TGH6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Guanosine diphosphatase (EC 3.6.1.42).
GN GDA1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RA Herrero A.B.;
RT "The Golgi GDPase of the fungal pathogen Candida albicans Affects
RT morphogenesis, glycosylation and cell wall properties.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ421721; CAD18870.1; -.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW Hydrolase.
SQ SEQUENCE 599 AA; 65952 MW; D5BC3A4F6E1B646A CRC64;

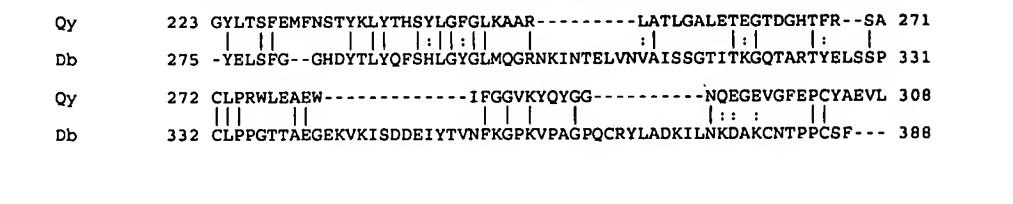
Query Match 22.9%; Score 515.5; DB 3; Length 599;
Best Local Similarity 32.6%; Pred. No. 8.5e-36;
Matches 143; Conservative 65; Mismatches 148; Indels 83; Gaps 15;





RESULT 14
Q9HEM6
ID Q9HEM6 PRELIMINARY; PRT; 522 AA.
AC Q9HEM6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Guanosine diphosphatase.
GN GDA1.
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MG1/2;
RX MEDLINE=21319012; PubMed=11425802;
RA Lopez-Avalos M.D., Uccelletti D., Abeijon C., Hirschberg C.B.;
RT "The UDPase activity of the Kluyveromyces lactis Golgi GDPase has a
RT role in uridine nucleotide sugar transport into Golgi vesicles.";
RL Glycobiology 11:413-422(2001).
DR EMBL; AJ401304; CAC21576.1; -.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
SQ SEQUENCE 522 AA; 56863 MW; A7A27902607A6732 CRC64;

Query Match 22.7%; Score 511; DB 3; Length 522;
Best Local Similarity 33.8%; Pred. No. 1.7e-35;
Matches 147; Conservative 59; Mismatches 149; Indels 80; Gaps 17;



```
Qy        309 RVVRGKLHQPEEV----QRGSFYAFSYYYDRAVDTDMIDYEKGGI---LKVEDFERKARE 361
                  :|||  |    :    | |||:|||         :  |:     ::: :  ||
Db        389 ----NGIHQPSLVHTFKETSDLYVFSYFYDRT--------QPLGLPLSFTLQELQDLART 436

Qy        362 VCDNLENFTS------GS-------PFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNI 408
              ||:  | : |      ||       |  |:||::  :||  |:        |:  | : |
Db        437 VCNGEEVWESVFSGIEGSLSELSKEPQWCLDLNFQVSLLHTGYDIPLQRELRTAKTIANN 496

Qy        409 ETGWALGATFHLLQS 423
              | || |||:  ||:|
Db        497 ELGWCLGASLPLLES 511


RESULT 15
Q9SPM6
ID   Q9SPM6      PRELIMINARY;      PRT;   455 AA.
AC   Q9SPM6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Nod factor binding lectin-nucleotide phosphohydrolase.
GN   LNP.
OS   Medicago sativa (Alfalfa).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX   NCBI_TaxID=3879;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Root;
RX   MEDLINE=99444909; PubMed=10517321;
RA   Roberts N.J., Brigham J., Wu B., Murphy J.B., Volpin H.,
RA   Phillips D.A., Etzler M.E.;
RT   "A Nod factor-binding lectin is a member of a distinct class of
RT   apyrases that may be unique to the legumes.";
RL   Mol. Gen. Genet. 262:261-267(1999).
DR   EMBL; AF156782; AAF00611.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Lectin.
SQ   SEQUENCE   455 AA;  49879 MW;  68122846D7EC261B CRC64;

  Query Match           22.6%;  Score 507.5;  DB 10;  Length 455;
  Best Local Similarity  34.0%;  Pred. No. 2.7e-35;
  Matches  143;  Conservative   65;  Mismatches  152;  Indels   61;  Gaps   16;

Qy         47 TLYGIMFDAGSTGTRIHVYTFVQKMPGQLPIL----EGEVFDSVKPGLSAFVDQPKQGAE 102
              | | ::|||||||||:||| | |     | :|    : |  | :||||||: | |:| |:
Db         43 TSYAVIFDAGSTGTRVHVYHFDQ----NLDLLHIGNDIEFVDKIKPGLSAYGDNPEQAAK 98

Qy        103 TVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIF-RKSPFLVPK 161
              ::  ||| |:| :|     |||: | |||||||     |: :|   : :|  :|   | :
Db         99 SLIPLLEEAEDVVPEDLHPKTPLRLGATAGLRLLNGDAAEKILQATRNMFSNRSTLNVQR 158

Qy        162 GSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTP 221
               :|||:||: ||   |||||:: | |    ::|| :|||| | |:|:    :|| :  |
Db        159 DAVSIIDGTQEGSYMWVTVNYVLGNLGKSFTKSVGVIDLGGGSVQMTYAVS-KKTAKNAP 217

Qy        222 R------GYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRSACLPR 275
              :      |:    :   | || |||| || :| |   | |  | |:      : |:
Db        218 KVADGEDPYIKKLVLKGKQYDLYVHSYLRFGKEATRAQVLNA--TNGS-----ANPCILP 270

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSF--------- 326
                  : : ||:|:      |   |: |   :|:|:  |:: |       :|
Db        271 GFNGTFTYSGVEYKAFSPSSGS-NFDDCKEIILKVL--KVNDPCPYPSCTFGGIWNGGGG 327

Qy        327 ---------YAFSYYYDRAVDTDMIDYEK-GGILKVEDFERKAREVCD-NLENFTS---- 371
                      ||:|    | |  |::  |   ||   ||| :|:  |  | |:  |
Db        328 SGQKKLFVTSAFAYL---AEDVGMVEPNKPNSILHPVDFEIEAKRACALNFEDVKSTYPR 384

Qy        372 ---GSPFLCMDLSYITALLKDGFGFA---DSTVLQLTKKVNN-IETGWALGATFHLLQS 423
                                                |::|||| |   ||  |||     : || :  :  |: :|  | ||     : :
Db        385 LTDAKRPYVCMDLLYQHVLLVHGFGLGPRKEITVGEGIQYQNSVVEAAWPLGTAVEAISA 444

Qy        424 L 424
              |
Db        445 L 445


Search completed: January 29, 2004, 10:05:30
Job time : 36 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:58:54 ; Search time 16.5 Seconds
(without alignments)
2494.557 Million cell updates/sec

Title: US-10-091-085-3
Perfect score: 2250
Sequence: 1 MATSWGTVFFMLVVSCVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	616.5	27.4	479	2	T23508	hypothetical prote
2	538.5	23.9	556	2	T39109	probable guanosine
3	489	21.7	455	2	S48859	nucleoside triphos
4	483.5	21.5	518	2	A40732	guanosine-diphosph
5	446.5	19.8	454	2	JC4616	apyrase (EC 3.6.1.
6	425.5	18.9	516	2	G84442	probable nucleosid
7	397	17.6	483	2	D86276	hypothetical prote
8	387.5	17.2	485	2	T34147	hypothetical prote
9	372.5	16.6	557	2	T16696	hypothetical prote
10	372	16.5	630	2	S50463	hypothetical prote
11	351	15.6	510	2	I56242	lymphoid cell acti
12	347	15.4	572	2	T40856	probable nucleotid
13	329.5	14.6	405	2	E86276	hypothetical prote
14	274	12.2	1052	2	T04439	hypothetical prote
15	248	11.0	508	2	C86276	7A19.33 protein -
16	138	6.1	628	2	A55421	nucleoside-triphos
17	110	4.9	497	1	JC2192	subtilisin-like pr
18	106.5	4.7	553	1	G1BPSV	gene 1 protein - s
19	105.5	4.7	774	2	T14555	DNA polymerase hom
20	104.5	4.6	684	2	T47694	probable serine/th
21	104.5	4.6	4427	2	PN0637	polyketide synthas
22	102.5	4.6	1019	2	T40813	probable cell divi
23	101.5	4.5	535	2	F97910	glucan 1,6-alpha-g
24	101.5	4.5	1151	2	H71347	hypothetical prote
25	99.5	4.4	535	2	C95040	glucan 1,6-alpha-g
26	99.5	4.4	548	1	HYBSS	bacillolysin (EC 3
27	97.5	4.3	604	2	F89453	protein F35H12.4 [
28	97.5	4.3	654	2	AG3522	iron-regulated out
29	97	4.3	1220	2	T18291	patched protein -
30	96	4.3	711	2	S66749	hypothetical prote
31	95.5	4.2	2076	2	S15999	fatty-acyl-CoA syn
32	95	4.2	652	1	JC2191	subtilisin-like pr
33	93	4.1	478	2	F90497	hypothetical prote
34	93	4.1	635	1	T37835	probable phosphoes
35	93	4.1	962	2	JC5571	subtilisin-like pr
36	93	4.1	969	1	A39490	subtilisin-like pr
37	93	4.1	975	2	JC5570	subtilisin-like pr
38	92	4.1	290	2	A87312	coproporphyrinogen
39	92	4.1	528	2	S13641	H+-transporting tw
40	91	4.0	417	2	H70174	glycine hydroxymet
41	91	4.0	477	2	JS0597	t-plasminogen acti
42	91	4.0	500	2	D83984	gluconate kinase g
43	91	4.0	726	2	AB0122	probable ferric si
44	90.5	4.0	301	2	E71482	phosphatidylserine
45	90.5	4.0	377	2	C82113	succinyl-diaminopi

ALIGNMENTS

RESULT 1
T23508
hypothetical protein K08H10.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C;Accession: T23508
R;Gardner, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19750
A;Accession: T23508
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-479 <WIL>
A;Cross-references: EMBL:Z83113; PIDN:CAB05544.1; GSPDB:GN00023; CESP:K08H10.4
A;Experimental source: clone K08H10
C;Genetics:
A;Gene: CESP:K08H10.4
A;Map position: 5
A;Introns: 36/3; 83/3; 189/1; 300/2; 412/3
C;Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 27.4%; Score 616.5; DB 2; Length 479;
Best Local Similarity 34.3%; Pred. No. 1.3e-43;
Matches 150; Conservative 81; Mismatches 167; Indels 39; Gaps 16;

```
Qy          9 FFMLVVS--CVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYT 66
              | :|::|   : | |: : |  |  |  : :    |       : |: ||||||||:|:|
Db          6 FSILLISFFSLLSVVTTKTQY-WCHGDGVLN----NQHTCRFFTIVIDAGSTGTRLHLYK 60

Qy         67 FVQK-------MPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSH 119
              |:          || :   :| |:|  ||||||:|   |    |:::: ||: |:  :|
Db         61 FIHDPAIASHGMPFK---VEKEIFQEVKPGLSSFAKSPSSAADSLEPLLQRARKEVPHFM 117

Qy        120 WKKTPVVLKATAGLRLLPEHKAKALLFEVKE-IFRKSPFLVPKGSVSIMDGSDEGILAWV 178
             |:|||: ||||||||||    | :|  |:| ||    |     :|::| |||||: :|
Db        118 WEKTPITLKATAGLRLLPGDMADDILESVEERIFNSGFFAAFPDAVNVMPGSDEGVYSWF 177

Qy        179 TVNFLTGQLH------GHR---QETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFE 229
             |:| |   |        ||:     :|   |||| |||:|: |  |    :   ||    :
Db        178 TLNILLETLFTDEPTVGHKPAAHRSVAAFDLGGGSTQLTYWPNNEAVFSEHV-GYERDID 236

Qy        230 MFNSTYKLYTHSYLGFGLKAARLATLGALETEG--TDGHTFRSACLPRWLE-AEWIFGGV 286
             |    :|:|||:|| || ||||  |  |||:       |   ::|:|   :  || :  :
Db        237 FFGHHIRLFTHSFLGNGLIAARLNIL-QLETDNEIESTHQLITSCMPEGYQLTEWEY-AL 294

Qy        287 KYQYGGNQEGEVGFEPCYAEVLRVVR-GKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYE 345
             |: :  |       || ||       |   ::    |::    | |||::|||::: ::
Db        295 KF-WNINGSSSHSFESCYGTTKNFVESSEIMHLRELKGSPVYLFSYFFDRALNSGLVKGN 353

Qy        346 KGGILKVEDFERKAREVC----DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQL 401
             :|| :::  |:  |   |      ::: :   |: |:||:|| :||:||: | |:  | |
Db        354 EGGKIELRQFKEAAEIACRREKTEIDDGSHWMPWQCLDLTYIYSLLRDGYQFEDNQPLVL 413
```

```
Qy        402 TKKVNNIETGWALGATF 418
              ||:  :|  |  |  |
Db        414 AKKIKGMEVSWGQGLAF 430
```

RESULT 2
T39109
probable guanosine-diphosphatase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T39109
R;Barrell, B.G.; Rajandream, M.A.; Quail, M.; Seegar, K.; Harris, D.
submitted to the EMBL Data Library, October 1999
A;Reference number: Z21828
A;Accession: T39109
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-556 <BAR>
A;Cross-references: EMBL:AL121741; PIDN:CAB57338.1; GSPDB:GN00066; SPDB:SPAC824.08
A;Experimental source: strain 972h-; cosmid c824
C;Genetics:
A;Gene: SPDB:SPAC824.08
A;Map position: 1

  Query Match 23.9%; Score 538.5; DB 2; Length 556;
  Best Local Similarity 34.4%; Pred. No. 5.7e-37;
  Matches 144; Conservative 61; Mismatches 157; Indels 57; Gaps 13;

```
Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLL 108
              | :| ||||||:|:||| |    |  | || | |  ::||||:|   |:  | ::  ||
Db        134 YVLMIDAGSTGSRVHVYQFNNCNPS--PKLEEEFFKMIEPGLSSFAGDPEGAAASLDPLL 191

Qy        109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKS-PFLVPKGSVSIM 167
              : | :::|  : : :|: :|||||||| | :|||:|  |::      || : |  |||:
Db        192 DYAMENVPEEYRRCSPIAVKATAGLRLTGESEAKAILKSVRQHLENDYPFPIVKDGVSIL 251

Qy        168 DGSDEGILAWVTVNFLTGQLHGH-RQETVGTLDLGGASTQITFLPQFEKTLEQTPRG--- 223
              :|| ||| ||:|:|:| | | |     ||  :||||||||: | |:|    |   |
Db        252 EGSMEGIYAWITINYLLGTLGGKATHSTVAVMDLGGASTQLVFEPRFASDGESLVDGDHK 311

Qy        224 YLTSFEMFNSTYKLYTHSYLGFGLKAARL----------ATLGALETEGTDGHTFRSAC 272
              |:  :      |:|| ||:||:||| ||           |   :||  | |  :    |
Db        312 YVLDYN--GEQYELYQHSHLGYGLKEARKLIHKFVLNNAEALKESLELLG-DSTSIIHPC 368

Qy        273 L----------PRWLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVR----GKLHQP 318
              |          :  :| :| |    :   |  : :  | :    ||      :|||
Db        369 LHLNASLTHPDSKSEASEVVFVGPSLAHLSLQCRGIAEKALYKDKNCPVRPCSFNGVHQP 428

Qy        319 ---EEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAREVC---------- 363
                 |      |  ||:|||     ||         :|| :  |  ||
Db        429 KFTETFTDSPIYLISYFYDR-----MISLGMPSTFTIEDMKYLANSVCSGPTYWQDAFSL 483

Qy        364 -DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHLL 421
               |:    |  |:||:|: :||  |:   ::  |   ||::| | || |||:  :|
Db        484 TDALKELKE-EPEWCLDLNYMISLLSVGYEIPNNRQLHTAKKIDNKELGWCLGASLSML 541
```

RESULT 3
S48859
nucleoside triphosphatase precursor, chromatin-associated - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 29-Sep-1999
C;Accession: S65147; S48859
R;Hsieh, H.L.; Tong, C.G.; Thomas, C.; Roux, S.J.
Plant Mol. Biol. 30, 135-147, 1996
A;Title: Light-modulated abundance of an mRNA encoding a calmodulin-regulated, chromatin
A;Reference number: S65141; MUID:96197404; PMID:8616230
A;Accession: S65147
A;Molecule type: mRNA
A;Residues: 1-455 <HS2>
A;Cross-references: EMBL:Z32743; NID:g563611; PIDN:CAA83655.1; PID:g563612
C;Superfamily: nucleoside triphosphatase chromatin-associated
C;Keywords: nucleus

  Query Match 21.7%; Score 489; DB 2; Length 455;
  Best Local Similarity 33.7%; Pred. No. 6e-33;
  Matches 140; Conservative 62; Mismatches 160; Indels 54; Gaps 14;

```
Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEG-EVFDSVKPGLSAFVDQPKQGAETVQGL 107
              | ::|||||||:|||| | | :   | | :| | :: : ||||:: : |:| |:::  |
Db         44 YAVVFDAGSTGSRIHVYHFNQNL-DLLHIGKGVEYYNKITPGLSSYANNPEQAAKSLIPL 102

Qy        108 LEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIF-RKSPFLVPKGSVSI 166
              || |:| :|     |||| | ||||||||    :: :|  |:::   :| | |  .:|||
Db        103 LEQAEDVVPDDLQPKTPVRLGATAGLRLLNGDASEKILQSVRDMLSNRSTFNVQPDAVSI 162

Qy        167 MDGSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPR---- 222
              :||: ||   |||||:  | |     :||| :|||| | |: :    :|| :   |:
Db        163 IDGTQEGSYLWVTVNYALGNLGKKYTKTVGVIDLGGGSVQMAYAVS-KKTAKNAPKVADG 221

Qy        223 --GYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRS--ACLPRWLE 278
                |:   .:    | || |||| || :|:|   |            | ||   ||
Db        222 DDPYIKKVVLKGIPYDLYVHSYLHFGREASRAEILKL---------TPRSPNPCLLAGFN 272

Qy        279 AEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSF------------ 326
                : : | :::      |   |  |   : : :  ||: |   |  :|
Db        273 GIYTYSGEEFKATAYTSG-ANFNKCKNTIRKAL--KLNYPCPYQNCTFGGIWNGGGGNGQ 329

Qy        327 ---YAFSYYYDRAVDTDMIDYEKGG-ILKVEDFERKAREVCD-NLENFTSGSPFL----- 376
                :| | ::    || |:|      ||:  | | ||:| |  | | |:  |  |||
Db        330 KNLFASSSFFYLPEDTGMVDASTPNFILRPVDIETKAKEACALNFEDAKSTYPFLDKKNV 389

Qy        377 ----CMDLSYITALLKDGFGFADSTVLQLTKKVNN----IETGWALGATFHLLQSL 424
                   |||| |   || ||||      :   |::      :|  | ||    : :|
Db        390 ASYVCMDLIYQYVLLVDGFGLDPLQKITSGKEIEYQDAIVEAAWPLGNAVEAISAL 445
```

RESULT 4
A40732
guanosine-diphosphatase (EC 3.6.1.42) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YEL042w
C;Species: Saccharomyces cerevisiae
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 02-Sep-2000
C;Accession: A40732; B40732; S30837; S50502
R;Abeijon, C.; Yanagisawa, K.; Mandon, E.C.; Hausler, A.; Moremen, K.; Hirschberg, C.B.;
J. Cell Biol. 122, 307-323, 1993
A;Title: Guanosine diphosphatase is required for protein and sphingolipid glycosylation
A;Reference number: A40732; MUID:93308137; PMID:8391537
A;Accession: A40732
A;Molecule type: DNA
A;Residues: 1-518 <ABE>
A;Cross-references: EMBL:L19560; NID:g349392; PIDN:AAA34656.1; PID:g349393
A;Note: sequence extracted from NCBI backbone (NCBIN:134708, NCBIP:134711)
A;Accession: B40732
A;Molecule type: protein
A;Residues: 125-144;238-257;276-281;366-374;399-412 <AB2>
R;Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor, I
submitted to the EMBL Data Library, February 1993
A;Reference number: S30812
A;Accession: S30837
A;Molecule type: DNA
A;Residues: 1-518 <MUL>
A;Cross-references: GB:U18779; EMBL:L10830; NID:g603625; PIDN:AAB65000.1; PID:g603637
R;Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A;Description: The sequence of S. cerevisiae cosmids 8199, 8334, and 9871.
A;Reference number: S50491
A;Accession: S50502
A;Molecule type: DNA
A;Residues: 1-518 <DIE>
A;Cross-references: EMBL:U18779; NID:g603625; PIDN:AAB65000.1; PID:g603637; MIPS:YEL042w
C;Genetics:
A;Gene: SGD:GDA1
A;Cross-references: SGD:S0000768; MIPS:YEL042w
A;Map position: 5L
```
```

C;Function:
A;Description: hydrolase
C;Superfamily: nucleoside triphosphatase chromatin-associated
C;Keywords: glycoprotein; Golgi apparatus; hydrolase; transmembrane protein
F;10-24/Domain: transmembrane #status predicted <TMM>
F;41,280,335/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 21.5%; Score 483.5; DB 2; Length 518;
Best Local Similarity 32.8%; Pred. No. 2.1e-32;
Matches 142; Conservative 56; Mismatches 160; Indels 75; Gaps 15;

```
Qy          49 YGIMFDAGSTGTRIHVYTF-VQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGL 107
               |  || |||||||:|:|:| | |    |   | |  | || ::||||:|       | ::  |
Db          93 YVIMIDAGSTGSRVHIYKFDVCTSP---PTLLDEKFDMLEPGLSSFDTDSVGAANSLDPL 149

Qy         108 LEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKS-PFLVPKG-SVS 165
               |:|| : :|      ||| :||||||||| : |:  :|  |::   |  || | :|  ||
Db         150 LKVAMNYVPIKARSCTPVAVKATAGLRLLGDAKSSKILSAVRDHLEKDYPFPVVEGDGVS 209

Qy         166 IMDGSDEGILAWVTVNFLTGQL--HGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRG 223
               || | :||: ||:| |:| | :  :| :  |    ||||  |||| | | |    :
Db         210 IMGGDEEGVFAWITTNYLLGNIGANGPKLPTAAVFDLGGGSTQIVFEPTFPINEKMVDGE 269

Qy         224 YLTSFEMFNSTYKLYTHSYLGFGLKAAR----------LATLGALETEGTDGHTFRSAC 272
               :    :  : | ||  |:||:|||  |          |    |: : |  |  | |
Db         270 HKFDLKFGDENYTLYQFSHLGYGLKEGRNKVNSVLVENALKDGKILKGDNTKTHQLSSPC 329

Qy         273 LPRWLEA--EWI----------FGGVKYQYGG----------NQEGEVGFEPCYAEVLR 309
               ||  : |  | :              | |     |          |:: :   ||
Db         330 LPPKVNATNEKVTLESKETYTIDFIGPDEPSGAQCRFLTDEILNKDAQCQSPPCSF---- 385

Qy         310 VVRGKLHQPEEV----QRGSFYAFSYYYDRAVDTDM-IDYEKGGILKVEDFERKAREVC- 363
                    :|||  |    :    | |||:|||     | : :       : :    || ||
Db         386 ---NGVHQPSLVRTFKESNDIYIFSYFYDRTRPLGMPLSF------TLNELNDLARIVCK 436

Qy         364 ------------DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIET 410
                          :|:  | | | |:|||:  :||  |:       |:  ||: | |
Db         437 GEETWNSVFSGIAGSLDELESDSHF-CLDLSFQVSLLHTGYDIPLQRELRTGKKIANKEI 495

Qy         411 GWALGATFHLLQS 423
               || |||:  ||::
Db         496 GWCLGASLPLLKA 508
```

RESULT 5
JC4616
apyrase (EC 3.6.1.5) precursor - potato
N;Alternate names: adenylpyrophosphatase; ATP-diphosphohydrolase
C;Species: Solanum tuberosum (potato)
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 28-May-1999
C;Accession: JC4616; PC4147
R;Handa, M.; Guidotti, G.
Biochem. Biophys. Res. Commun. 218, 916-923, 1996
A;Title: Purification and cloning of a soluble ATP-diphosphohydrolase (Apyrase) from pot
A;Reference number: JC4616; MUID:96158985; PMID:8579614
A;Accession: JC4616
A;Molecule type: mRNA
A;Residues: 1-454 <HAN>
A;Cross-references: GB:U58597; NID:g1381632; PIDN:AAB02720.1; PID:g1381633
A;Accession: PC4147
A;Molecule type: protein
A;Residues: 59-95;96-131;132-160;236-253;332-345 <HA2>
A;Experimental source: tubers
A;Note: The authors translated the codon GCA for residue 215 as Gly
C;Comment: This enzyme belongs to a family of E-type ATPases, and it catalyzes the hydro
enzyme has nucleotide substrate specificity, divalent cation requirement, and insensiti
olved in starch synthesis.
C;Genetics:
A;Gene: rrop1
C;Superfamily: nucleoside triphosphatase chromatin-associated
C;Keywords: glycoprotein; hydrolase; transmembrane protein
F;1-30/Domain: signal sequence #status predicted <SIG>
F;8-25/Domain: transmembrane #status predicted <TMM>
F;31-454/Product: ATP-diphosphohydrolase #status predicted <MAT>
F;44-65/Region: actin-heat shock protein 70-hexokinase beta-phosphate binding
F;44-65/Region: nucleotide binding #status predicted
F;192-212/Region: actin-heat shock protein 70-hexokinase gamma-phosphate binding
F;192-212/Region: nucleotide binding #status predicted
F;390-410,427-446/Region: hydrophobic carboxyl end
F;151,262/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.8%; Score 446.5; DB 2; Length 454;
Best Local Similarity 28.8%; Pred. No. 2.2e-29;
Matches 131; Conservative 78; Mismatches 177; Indels 69; Gaps 16;

```
Qy          25 NQQTWFEGIFLS-------SMCPINVSA-----------STLYGIMFDAGSTGTRIHVY 65
               || : |  | |:       |:   ||:|           |  | ::|||||||:|:||:
Db           3 NQNSHFIFIILAIFLVLPLSLLSKNVNAQIPLRRHLLSHESEHYAVIFDAGSTGSRVHVF 62

Qy          66 TFVQKMPGQLPILEG-EVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTP 124
                 | :|: | |||    | | : :||||:: : ||  | ::: ||: |:  :|:    :||
Db          63 RFDEKL-GLLPIGNNIEYFMATEPGLSSYAEDPKAAANSLEPLLDGAEGVVPQELQSETP 121

Qy         125 VVLKATAGLRLLPEHKAKALLFEVKEIFR-KSPFLVPKGSVSIMDGSDEGILAWVTVNFL 183
               : | ||||||:|    |: :|  |: : : :| |       |:|:||: ||   |  :|:|
Db         122 LELGATAGLRMLKGDAAEKILQAVRNLVKNQSTFHSKDQWVTILDGTQEGSYMWAAINYL 181

Qy         184 TGQLHGHRQETVGTLDLGGASTQITFL---PQFEKTLE-QTPRGYLTSFEMFNSTYKLYT 239
               | |    : |  |:|||| | |: :     || |  : :    |:    : :  | ||
Db         182 LGNLGKDYKSTTATIDLGGGSVQMAYAISNEQFAKAPQNEDGEPYVQQKHLMSKDYNLYV 241

Qy         240 HSYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVG 299
               |||| :|  | |      |  |        : |     :  : :|||  |:    ::|
Db         242 HSYLNYGQLAGRAEIFKASRNES-------NPCALEGCDGYYSYGGVDYKVKAPKKGS-- 292

Qy         300 FEPCYAEVLRVVRG--KLHQPEEVQRGSF--------------YAFSYYYDRAVDTDMI 342
                    :   |: |   |::   ::  :|                :| |::||      ::
Db         293 ---SWKRCRRLTRHALKINAKCNIEECTFNGVWNGGGGDGQKNIHASSFFYDIGAQVGIV 349

Qy         343 DYE-KGGILKVEDFERKAREVCD-NLENFTS--------GSPFLCMDLSYITALLKDGFG 392
               | :   : |   :   |:  |  |: : |          |:|||| |   || ||||
Db         350 DTKFPSALAKPIQYLNAAKVACQTNVADIKSIFPKTQDRNIPYLCMDLIYEYTLLVDGFG 409

Qy         393 F---ADSTVLQLTKKVNN-IETGWALGATFHLLQS 423
                    : ||:   :  |  :   | ||    |: |
Db         410 LNPHKEITVIHDVQYKNYLVGAAWPLGCAIDLVSS 444
```

RESULT 6
G84442
probable nucleoside triphosphatase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 18-Nov-2002
C;Accession: G84442
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84442
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-516 <STO>
A;Cross-references: GB:AE002093; NID:g3461821; PIDN:AAC32915.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g02970
A;Map position: 2
C;Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 18.9%; Score 425.5; DB 2; Length 516;
Best Local Similarity 30.6%; Pred. No. 1.5e-27;
Matches 127; Conservative 61; Mismatches 154; Indels 73; Gaps 13;

```
Qy          42 NVSASTLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILE--GEVFDSVK--PGLSAFVDQP 97
               |:  |  | :: | |||||||||: :  ::    |: |  |  : |:|  |||||| | |
```

```
Db         61 NLRGSLRYSVVIDGGSTGTRIHVFGY--RIESGKPVFEFRGANYASLKLHPGLSAFADDP 118

Qy         98 KQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPF 157

Db        119 DGASVSLTELVEFAKGRVPKGMWIETEVRLMATAGMRLLELPVQEKILGVARRVLKSSGF 178

Qy        158 LVPKGSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFL------P 211

Db        179 LFRDEWASVISGSDEGVYAWVVANFALGSLGGDPLKTTGIVELGGASAQVTFVSSEPMPP 238

Qy        212 QFEKTLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGAL-------ETEGTD 264

Db        239 EFSRTI---------SFG--NVTYNLYSHSFLHFGQNAAHDKLWGSLLSRDHNSAVEPTR 287

Qy        265 GHTFRSACLPR-------------WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVV 311

Db        288 EKIFTDPCAPKGYNLDANTQKHLSGLLAEESRLSDSFQAGGN------YSQCRSAALTIL 341

Qy        312 RGKLHQPEEVQRGSFYAFSYYYDRAVDTDMI---------DYEKGGILKVEDFERKAREV 362

Db        342 QDGNGRILIIIAGFSFLFFGLGEKAWLSNMISAGERFCGEDWSK---LRVKDPSLHEEDL 398

Qy        363 CDNLENFTSGSPFLCMDLSYITALLKDGFGF-ADSTVLQLTKKVNNIETGWALGA 416

Db        399 LR-----------YCFSSAYIVSLLHDTLGIPLDDERIGYANQAGDIPLDWALGA 442


RESULT 7
D86276
hypothetical protein F7A19.34 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 18-Nov-2002
C;Accession: D86276
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D86276
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-483 <STO>
A;Cross-references: GB:AE005172; NID:g5080801; PIDN:AAD39311.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: nucleoside triphosphatase chromatin-associated

  Query Match             17.6%;  Score 397;  DB 2;  Length 483;
  Best Local Similarity   27.6%;  Pred. No. 3.3e-25;
  Matches  120;  Conservative   71;  Mismatches  162;  Indels   82;  Gaps   17;

Qy         32 GIFLSSMCPINVSASTLYGIMFDAGSTGTRIHV--YTFVQKMPGQLPILEGEVFD----- 84

Db         59 GSLLSRRCKLR------YSVLIDAGSSGTRVHVFGYWFESGKP---------VFDFGEKH 103

Qy         85 ----SVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHK 140

Db        104 YANLKLTPGLSSYADNPEGASVSVTKLVEFAKQRIPKRMFRRSDIRLMATAGMRLLEVPV 163

Qy        141 AKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETVGTLDL 200

Db        164 QEQILEVTRRVLRSSGFMFRDEWANVISGSDEGIYSWITANYALGSLGTDPLETTGIVEL 223

Qy        201 GGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALET 260

Db        224 GGASAQVTFV-----SSEHVPPEYSRTIAYGNISYTIYSHSFLDYGKDAALKKLLEKLQN 278

Qy        261 EG---TDGHTFRSACLPRWLEAEWIF--GGVKYQYG---------GNQEGEVGFEPCYAE 306

Db        279 SANSTVDG-VVEDPCTPK----GYIYDTNSKNYSSGFLADESKLKGSLQAAGNFSKCRSA 333

Qy        307 VLRVVR-------------GKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKV- 352

Db        334 TFALLKEGKENCLYEHCSIGSTFTPD--LQGSFLATASFYYTAKFFEL--EEKGWLSELI 389

Qy        353 --------EDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFA--DSTVLQLT 402

Db        390 PAGKRYCGEEWSKLILEYPTTDEEYLRG---YCFSAAYTISMLHDSLGIALDDESITYAS 446

Qy        403 KK-VNNIETGWALGA 416

Db        447 KAGEKHIPLDWALGA 461


RESULT 8
T34147
hypothetical protein C33H5.14 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Mar-2000
C;Accession: T34147
R;Bradshaw, H.; Stellyes, L.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C33H5.
A;Reference number: Z21482
A;Accession: T34147
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-485 <BRA>
A;Cross-references: EMBL:U41007; PIDN:AAA82272.1; CESP:C33H5.14
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:C33H5.14
A;Introns: 22/1; 83/1; 120/2; 167/1; 269/3; 399/3
C;Superfamily: nucleoside triphosphatase chromatin-associated

  Query Match             17.2%;  Score 387.5;  DB 2;  Length 485;
  Best Local Similarity   27.7%;  Pred. No. 2.1e-24;
  Matches  120;  Conservative   84;  Mismatches  152;  Indels   77;  Gaps   18;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDS------VKPGLSAFVDQPKQGAE 102

Db         25 YGVICDAGSSGTRLFVYT-LKPLSGGLTNIDTLIHESEPVVKKVTPGLSSFGDKPEQVVE 83

Qy        103 TVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFL-VPK 161

Db         84 YLTPLLRFAEEHIPYEQLGETDLLIFATAGMRLLPEAQKDAIIKNLQNGLKSVTALRVSD 143

Qy        162 GSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEK------ 215

Db        144 SNIRIIDGAWEGIYSWIAVNYILGRFDKENDSKVGMIDMGGASVQIAFEIANEKESYNGG 203

Qy        216 TLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRSACLPR 275

Db        204 NVYEINLGSIETNEDYK--YKIYSTTFLGYGANEGLKKYENSLVKSGNS----NDSCSPR 257

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPE---------------- 319

Db        258 GLNR--LIG----EFTVNGTGE--WDVCLAQVSSLI-GDKAQPSCPNPTCFLRNVIAPSV 308

Qy        320 EVQRGSFYAFS-YYYDRAVDTDMIDYEKGGILKVEDFERKAREVC----DNLENFTSGSP 374

Db        309 NLSTVQLYGFSEYWYTTS------NFGSGGEYHYQKFTDEVRKYCQKDWNDIQDGFKRNE 362

Qy        375 F----------LCMDLSYITALLKDGFGFADST--VLQLTKKVNNIETGWALGA------ 416

Db        363 FPNADIERLGTNCFKAAWVTSVLHDGFN-VDKTKHLFQSVLKIAGEEMQWALGAMLYHSK 421

Qy        417 --TFHLLQSLGIS 427

Db        422 DLKFNLLEQLEVA 434
```

RESULT 9
T16696
hypothetical protein R07E4.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T16696
R;Miller, N.
submitted to the EMBL Data Library, October 1995
A;Description: The sequence of C. elegans cosmid R07E4.
A;Reference number: Z18561
A;Accession: T16696
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-557 <MIL>
A;Cross-references: EMBL:U39652; NID:g1049390; PID:g1049394; PIDN:AAA80403.1; CESP:R07E4
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:R07E4.4
A;Introns: 39/1; 67/1; 145/2; 192/1; 244/3; 317/3; 451/3; 488/2

Query Match 16.6%; Score 372.5; DB 2; Length 557;
Best Local Similarity 28.3%; Pred. No. 4.6e-23;
Matches 126; Conservative 69; Mismatches 160; Indels 91; Gaps 21;

```
Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDS------VKPGLSAFVDQPKQGAE 102
              ||:: |||||||: || ::     :|  :|  ::|:       : |||| |  :| | ||
Db         44 YGVICDAGSTGTRLFVYNWISTSDSELIQIEPVIYDNKPVMKKISPGLSTFGTKPAQAAE 103

Qy        103 TVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEH-----KAKALLFEVKEIFRKSPF 157
               :: |:|:|:  ||      ||| : ||||:||:|:      : :|:|   | :  | |
Db        104 YLRPLMELAERHIPEEKRPYTPVFIFATAGMRLIPDEYVLIGQKEAVL---KNLRNKLPK 160

Qy        158 L----VPKGSVSIMDGSDEGILAWVTVNFLTGQLH-----------GH-RQETVGTLDL 200
              :    | |  : |::|  ||| :|: ||:  |: :             | ||:||| :|:
Db        161 ITSMQVLKEHIRIIEGKWEGIYSWIAVNYALGKFNKTATLDFPGTSPAHARQKTVGMIDM 220

Qy        201 GGASTQITF-LPQFEK----TLEQTPRGYLTSFEMFNSTYKLYTHSYLGF----GLKAAR 251
              |||| || | ||  :      :|    |      :|   |||:  ::||:    |::
Db        221 GGASAQIAFELPDTDSFSSINVENINLGCREDDSLFK--YKLFVTTFLGYGVNEGIRKYE 278

Qy        252 LATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGN--QEG------------ 296
                |  |: :  :|   :  |:|  |        |  : | |  : |
Db        279 HMLLSKLKDQ--NGTVIQDDCMPLNLHKT-----VTLENGENFVRRGTGNWNTCSNEVKK 331

Qy        297 ----EVGFEPCYAEVLRVVRGKLHQPE-EVQRGSFYAFSYYYDRAVDTDMIDYEKGGILK 351
                  |  | | ||  :  | :  |   :      | || |:    |        ||
Db        332 LLNPESSSEVCKAEAAKCYFGAVPAPSIPLSNIEMYGFSEYWYSTHDV----LGLGGQYD 387

Qy        352 VEDFERKAREVCD------NLENFTSGSP--------FLCMDLSYITALLKDGFGFADST 397
               |:  :| :: |          |:     |          |   ::||::| |||   | |
Db        388 AENIAKKTQQYCSKRWSTIQAESKKQLYPRADEERLRTQCFKSAWITSVLHDGFS-VDKT 446

Qy        398 --VLQLTKKVNNIETGWALGA-TFHL 420
                  |    :   |  |||||  :|:
Db        447 HNKFQSVSTIAGQEVQWALGAMIYHM 472
```

RESULT 10
S50463
hypothetical protein YER005w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 19-Apr-2002
C;Accession: S50463
R;Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A;Description: The sequence of S. cerevisiae cosmids 9537, 9581, 9495, 9867, and lambda
A;Reference number: S50433
A;Accession: S50463
A;Molecule type: DNA
A;Residues: 1-630 <DIE>
A;Cross-references: EMBL:U18778; NID:g603592; PID:g603597; GSPDB:GN00005; MIPS:YER005w
C;Genetics:
A;Gene: SGD:YND1; MIPS:YER005w
A;Cross-references: SGD:S0000807
A;Map position: 5R

Query Match 16.5%; Score 372; DB 2; Length 630;
Best Local Similarity 26.2%; Pred. No. 6.1e-23;
Matches 117; Conservative 73; Mismatches 165; Indels 92; Gaps 17;

```
Qy         49 YGIMFDAGSTGTRIHVYTF------------VQKMPGQLPILEGE---VFDSVKPGLSA 92
              :||: ||||:|:||||: :                | :   :| :  |    | : ||||:
Db         10 FGIVIDAGSSGSRIHVFKWQDTESLLHATNQDSQSILQSVPHIHQEKDWTF-KLNPGLSS 68

Qy         93 FVDQPKQGAET-VQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEV-KE 150
              |  :|:   :: :: ||: ||: || |||   || ::||||:||||:    ::|  : :
Db         69 FEKKPQDAYKSHIKPLLDFAKNIIPESHWSSCPVFIQATAGMRLLPQDIQSSILDGLCQG 128

Qy        151 IFRKSPFLVP--KGSVSIMDGSDEGILAWVTVNFLTGQLHGHRQE-----TVGTLDLGGA 203
              :   : |||      : ::||  ||:  |: :|:| |  : :  |      | | :|:|||
Db        129 LKHPAEFLVEDCSAQIQVIDGETEGLYGWLGLNYLYGHFNDYNPEVSDHFTFGFMDMGGA 188

Qy        204 STQITFLPQFEKTLE----------QTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARL 252
              |||| | |     :           ::  | |  :::| ||        :||||   ||
Db        189 STQIAFAPHDSGEIARHRDDIATIFLRSVNGDLQKWDVFVST-------WLGFGANQARR 241

Qy        253 ATLGAL-----------ETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGPE 301
                | |           | :           |:||    :: |    :   |:   |   :
Db        242 RYLAQLINTLPENTNDYENDDFSTRNLNDPCMPRGSSTDFEFKDTIFHIAGSGNYEQCTK 301

Qy        302 PCYAEVLR--------VVRGKLHQPE-EVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKV 352
                  | :|:          :  :| | :      |    | |:  | |     :: ||
Db        302 SIYPLLLKNMPCDDEPCLFNGVHAPRIDFANDKFIGTSEYWYTANDV----FKLGGEYNF 357

Qy        353 EDFERKAREVCDN-----LENFTSG------SPFL---CMDLSYITALLKDGFGFA---- 394
              : | :  || |::      | |   |          ||   |   :::  :| :||
Db        358 DKFSKSLREFCNSNWTQILANSDKGVYNSIPENFLKDACFKGNWVLNILHEGFDMPRIDV 417

Qy        395 ------DSTVLQLTKKVNNIETGWALG 415
                     |  : |  :||   |  | ||
Db        418 DAENVNDRPLFQSVEKVEERELSWTLG 444
```

RESULT 11
I56242
lymphoid cell activation antigen - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Nov-2002
C;Accession: I56242
R;Maliszewski, C.R.; Delespesse, G.J.; Schoenborn, M.A.; Armitage, R.J.; Fanslow, W.C.; N
J. Immunol. 153, 3574-3583, 1994
A;Title: The CD39 lymphoid cell activation antigen. Molecular cloning and structural char
A;Reference number: I56242; MUID:95015846; PMID:7930580
A;Accession: I56242
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-510 <RES>
A;Cross-references: GB:S73813; NID:g765255; PIDN:AAB32152.1; PID:g765256
C;Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 15.6%; Score 351; DB 2; Length 510;
Best Local Similarity 27.8%; Pred. No. 2.6e-21;
Matches 123; Conservative 71; Mismatches 175; Indels 74; Gaps 19;

```
Qy         32 GIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVK-PGL 90
              |:  :   | ||     |||: |||: | :::| :  :       ::       || ||:
Db         36 GLTQNKALPENVK----YGIVLDAGSSHTSLYIYKWPAEKENDTGVVHQVEECRVKGPGI 91

Qy         91 SAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLL---PEHKAKALLFE 147
              | || :  :    :  :| |:: ||||  ::||| | ||||:|||    |  |  :|
Db         92 SKFVQKVNEIGIYLTDCMERAREVIPRSQHQETPVYLGATAGMRLLRMESEELADRVLDV 151

Qy        148 VKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTVNFLTGQL-----------HGHRQETV 195
              |:      ||        |: | :||   |:|:|:| |:                : |||
```

```
Db        152 VERSLSNYPF--DFQGARIITGQEEGAYGWITINYLLGKFSQKTRWFSIVPYETNNQETF 209

Qy        196 GTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATL 255

Db        210 GALDLGGASTQVTFVPQ-NQTIE-SPDNAL-QFRLYGKDYNVYTHSFLCYGKDQALWQKL 266

Qy        256 GALETEGTDGHTFRSACL------------------PRWLEAEWIFGGVKYQYGGNQEGE 297

Db        267 -AKDIQVASNEILRDPCFHPGYKKVVNVSDLYKTPCTKRFEMTLPFQQFEIQGIGN---- 321

Qy        298 VGFEPCYAEVLRV------------VRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYE 345

Db        322 --YQQCHQSILELFNTSYCPYSQCAFNGIFLPPLQGDFGAFSAF-YFVMKFLN---LTSE 375

Qy        346 KGGILKVEDFERK-AREVCDNLENFTSG--SPFL---CMDLSYITALLKDGFGF-ADS-T 397

Db        376 KVSQEKVTEMMKKFCAQPWEEIKTSYAGVKEKYLSEYCFSGTYILSLLLQGYHFTADSWE 435

Qy        398 VLQLTKKVNNIETGWALGATFHL 420

Db        436 HIHFIGKIQGSDAGWTLGYMLNL 458


RESULT 12
T40856
probable nucleotide phosphatase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T40856
R;Ramsperger, U.; Pohl, T.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1999
A;Reference number: Z21952
A;Accession: T40856
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-572 <RAM>
A;Cross-references: EMBL:AL121783; PIDN:CAB57847.1; GSPDB:GN00068; SPDB:SPCC11E10.05c
A;Experimental source: strain 972h-; cosmid c11E10
C;Genetics:
A;Gene: SPDB:SPCC11E10.05c
A;Map position: 3

  Query Match           15.4%;  Score 347;  DB 2;  Length 572;
  Best Local Similarity   25.5%;  Pred. No. 6.6e-21;
  Matches  115;  Conservative   67;  Mismatches  157;  Indels  112;  Gaps   16;

Qy         49 YGIMFDAGSTGTRIHVYTF-------VQKMPGQLPILEGEVFD------SVKPGLSAFVD 95

Db          5 YGIFIDAGSSGSRLLIYSWDYDTDSSLSDKVKKLPLIETGIGDGGKWSLKVQPGISSFAN 64

Qy         96 QPKQ-GAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLL-PEHKAKALLFEVKEIFR 153

Db         65 NPKHVGKKHLKELLDFAAHAIPKDVHKETPVFLSATAGMRLLGVDAQNKILSHACRYIKK 124

Qy        154 KSPFLVP--KGSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLP 211

Db        125 NYDFDIPNCSNSIRVIDGKAEGMYGWLATNYLLKTLEEKDTSTVGFLDMGGASVQIAF-- 182

Qy        212 QFEKTLEQTPRGYLTSFEMFNST------------YKLYTHSYLGFGLKAARLATLGAL- 258

Db        183 -------ELPPSQLKNYKDSISTVHIGLQNGQQLEYPLFVTTWLGFGANEAYRRYLGLLI 235

Qy        259 ETE-GTDGHTFRSACLPRWLEAEWIFGGVKYQYGG------------NQEGEVGFEPCY 304

Db        236 ESENGKVGNTLSDPCSLR--GRTYDIDGIEFAGTGDLKQCLKLTYNLLNKDKPCSMDPCN 293

Qy        305 AEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAREVC- 363

Db        294 FDGISI------PPVDFANTEFVGVSEFWYTTNDV----FDMGGSYHFPNFYKKVDEYCG 343

Qy        364 --------------------DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTV---- 398

Db        344 TEWETMLSRLYNKELTPSTDENKLEK-------LCFKASWALNVLHEGFDVPKSNTSSND 396

Qy        399 --------------LQLTKKVNNIETGWALG 415

Db        397 AKDGLSVIPAYHSPFTSLEKIERTEVSWTLG 427


RESULT 13
E86276
hypothetical protein F14L17.1 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C;Accession: E86276
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E86276
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-405 <STO>
A;Cross-references: GB:AE005172; NID:g7262666; PIDN:AAF43924.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

  Query Match           14.6%;  Score 329.5;  DB 2;  Length 405;
  Best Local Similarity   26.3%;  Pred. No. 1.2e-19;
  Matches  110;  Conservative   63;  Mismatches  158;  Indels   87;  Gaps   15;

Qy         48 LYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGL 107

Db          3 VFGYWFESGKP-----VFDFGEEHYASL---------KLSPGLSSYADNPEGASVSVTKL 48

Qy        108 LEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIM 167

Db         49 VEFAKGRIPKGKLKKSDIRLMATAGMRLLDVPVQEQILDVTRRVLRSSGFKFQDEWATVI 108

Qy        168 DGSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTS 227

Db        109 SGTDEGIYAWVVANHALGSLGGDPLKTTGIVELGGASAQVTFVPS-----EHVPPEFSRT 163

Qy        228 FEMFNSTYKLYTHSYLGFGLKAARLATLGALETE---GTDGHTFRSACLPR--------- 275

Db        164 ISYGNVSYTIYSHSFLDFGQDAAEDKLLESLQNSVAASTGDGIVEDPCTPKGYIYDTHSQ 223

Qy        276 -----WLEAEWIF-GGVKYQYGGNQEGEVGFEPCYAEVLRVVR-GKLHQPEEVQRGSFYA 328

Db        224 KDSSGFLSEESKFKASLQVQAAGD------FTKCRSATLAMLQEGKEN------------ 265

Qy        329 FSYYYDRAVDTDMIDYEKGGILKVEDF----------------------ERKAREVCDNL 366

Db        266 -CAYKHCSIGSTFTPNIQGSFLATENFFHTSKFFGLGEKEWLSEMILAGKRFCGEEWSKL 324

Qy        367 -ENF-TSGSPFL---CMDLSYITALLKDGFGFA-DSTVLQLTKKV--NNIETGWALGA 416

Db        325 KEKYPTTKDKYLHRYCFSSAYIISMLHDSLGVALDDERIKYASKAGKENIPLDWALGA 382


RESULT 14
T04439
hypothetical protein T18B16.150 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
C;Accession: T04439
R;Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansorge, W.; Bancroft, I.; Mewes, H.W
submitted to the Protein Sequence Database, April 1998
A;Reference number: Z15359
```

```
A;Accession: T04439
A;Molecule type: DNA
A;Residues: 1-1052 <BEV>
A;Cross-references: EMBL:AL021687
A;Experimental source: cultivar Columbia; BAC clone T18B16
C;Genetics:
A;Map position: 4
A;Introns: 79/3; 118/3; 217/1; 295/3; 396/3; 531/2; 815/3
A;Note: T18B16.150

  Query Match             12.2%;  Score 274;  DB 2;  Length 1052;
  Best Local Similarity   24.9%;  Pred. No. 2.1e-14;
  Matches  107;  Conservative   70;  Mismatches  171;  Indels   82;  Gaps   19;

Qy         51 IMFDAGSTGTRIHVY--TFVQKMPGQLPIL------------EGEVFD--SVKPGLSAFV 94
              :::    |||| :||  :   .|    |||:            |  :|    :||   |
Db        525 LVYIVSITGTRAYVYQASINYKKDSSLPIVMKSLTEGISRKSRGRAYDRMETEPGFDKLV 584

Qy         95 DQPKQGAET-VQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFR 153
              :  : | :| :: |:: |:  ||::  : | : : ||||:| |    :  :|  |  |
Db        585 NN-RTGLKTAIKPLIQWAEKQIPKNAHRTTSLFVYATAGVRRLRPADSSWILGNVWSILA 643

Qy        154 KSPFLVPKGSVSIMDGSDEGILAWVTVNFLTGQLHG-HRQETVGTLDLGGASTQITFLPQ 212
              ||||   :  | |: |::|    |  :|: |  |    :: | | |||||:| |:||  :
Db        644 KSPFTCRREWVKIISGTEEAYFGWTALNYQTSMLGALPKKATFGALDLGGSSLQVTFENE 703

Qy        213 FEKTLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAA----------RLATL------- 255
               |:|  :|      :  : :  : |  :|  |:|| |          :|  :
Db        704 -ERTHNETN----LNLRIGSVNHHLSAYSLAGYGLNDAFDRSVVHLLKKLPNVNKSDLIE 758

Qy        256 GALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGFE------PCYAEVLR 309
              | ||          :  ||       ::|          | ::|: |         | : |
Db        759 GKLE--------MKHPCLNSGYNGQYICSQCASSVQGGKKGKSGVSIKLVGAPNWGECSA 810

Qy        310 VVR-GKLHQPEEVQR--GSFYAFSYYYDRAVDTDMIDYEKGGI---LKVEDFERKAREVC 363
              : :      |:   |  | ||| | ::       : |    :     ::|   | || |
Db        811 LAKNAPCALPDGYPRPHGQFYAVSGFF--------VVYRFFNLSAEASLDDVLEKGREFC 862

Qy        364 DN----LENFTSGSPFL---CMDLSYITALLKDGFGFADSTV------LQLTKKVNNIET 410
              |         |  ||:   |    || :||::|    |  :      :  |  | :|:
Db        863 DKAWQVARTSVSPQPFIEQYCFRAPYIVSLLREGLYITDKQIIIGSGSITWTLGVALLES 922

Qy        411 GWALGATFHL 420
              | || :|  |
Db        923 GKALSSTLGL 932


RESULT 15
C86276
7A19.33 protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 18-Nov-2002
C;Accession: C86276
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86276
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-508 <STO>
A;Cross-references: GB:AE005172; NID:g5080800; PIDN:AAD39310.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: nucleoside triphosphatase chromatin-associated

  Query Match             11.0%;  Score 248;  DB 2;  Length 508;
  Best Local Similarity   23.3%;  Pred. No. 1.1e-12;
  Matches  110;  Conservative   77;  Mismatches  159;  Indels  126;  Gaps   21;

Qy         12 LVVSCVCSAV----------SHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ::|:||  |:           | ||::                  |  | :: | ||:||
Db         35 VIVACVTIALGLLFIGYSILRSGRNRR----------------VSLHYSVIIDGGSSGT 77

Qy         61 RIHVYTFVQKMPGQLPILE--GEVFDSVK--PGLSAFVDQPKQGAETVQGLLEVAKDSIP 116
              |:||: :  ::    |: :   | : |:|  |||||: | |:  :|:|  |:| ||  :
Db         78 RVHVFGY--RIESGKPVFDFGEENYASLKLSPGLSAYADNPEGVSESVTELVEFAKKRVH 135

Qy        117 RSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDE--GI 174
              :  ||: : | ||||:|||     : :|   : : | | |              |  ||  :
Db        136 KGKLKKSDIRLMATAGMRLLELPVQEQILDVTRRVLRSSGF----------DFRDEWASV 185

Qy        175 LAWVTVNFLTGQLHGHRQETVGTLDLG------GASTQITFLPQFEKTLEQTPRGYLTSF 228
              :: :  ||          |: :    ||          ::||:     : | |  :  :
Db        186 ISEILENF---------QDLMKVYMLGLLLIMRSVRLEVTFV-----STELVPSEFSRTL 231

Qy        229 EMFNSTYKLYTHSYLGFGLKAARLATLGAL---ETEGTDGHTFRSACLPRWLEAEWI--- 282
              | :| ||:||:| ||  ||:     :|        |          |:|:    |
Db        232 AYGNVSYNLYSHSFLDFGQDAAQEKLSESLYNSAANSTGEGIVPDPCIPKGYILETNLQK 291

Qy        283 ----FGGVKYQYGGNQEGEVGFEPCYAEVLRVVR------------GKLHQPEEVQRGS 325
                  |  | ::    :   |  | :    :::              | :  |   :||
Db        292 DLPGFLADKGKFTATLQAAGNFSECRSAAFAMLQEEKGKCTYKRCSIGSIFTPN--LQGS 349

Qy        326 FYAFSYYY---------DRAVDTDMI---------DYEKGGILKVEDFERKAREVCDNLE 367
              | |   ::         ::   ::||         :: |   |||:    |     :||
Db        350 FLATENFFHTSKFFGLGEKEWLSEMILAGKRFCGEEWSK---LKVKYPTFKD----ENLL 402

Qy        368 NFTSGSPFLCMDLSYITALLKDGFGFA-DSTVLQLTKKV--NNIETGWALGA 416
              :        |   :|| ::| |  | | |   ::   |    :|  |||||
Db        403 RY-------CFSSAYIISMLHDSLGVALDDERIKYASKAGEEDIPLDWALGA 447


Search completed: January 29, 2004, 10:06:16
Job time : 19.5 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:35:09 ; Search time 9.5 Seconds
(without alignments)
2118.677 Million cell updates/sec

Title: US-10-091-085-3
Perfect score: 2250
Sequence: 1 MATSWGTVFFMLVVSCVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2250	100.0	428	1	ENP5_HUMAN	O75356 homo sapien
2	1994.5	88.6	427	1	ENP5_MOUSE	Q9wuz9 mus musculu
3	1933.5	85.9	469	1	ENP5_MESAU	Q9qyc8 mesocricetu
4	999	44.4	484	1	ENP6_HUMAN	O75354 homo sapien
5	986	43.8	455	1	ENP6_RAT	Q9er31 rattus norv
6	489	21.7	455	1	NTPA_PEA	P52914 pisum sativ
7	483.5	21.5	518	1	GDA1_YEAST	P32621 saccharomyc
8	446.5	19.8	454	1	APY_SOLTU	P80595 solanum tub
9	387.5	17.2	485	1	YY4E_CAEEL	Q18411 caenorhabdi
10	385	17.1	552	1	YBU4_CAEEL	Q21815 caenorhabdi
11	383	17.0	493	1	ENP1_CHICK	O93295 gallus gall
12	372	16.5	630	1	YND1_YEAST	P40009 saccharomyc
13	370	16.4	510	1	ENP1_MOUSE	P55772 mus musculu
14	365	16.2	513	1	ENP1_BOVIN	O18956 bos taurus
15	360.5	16.0	613	1	ENP4_MOUSE	Q9dbt4 mus musculu
16	357	15.9	616	1	ENP4_HUMAN	Q9y227 homo sapien
17	352	15.6	510	1	ENP1_PIG	Q9myu4 sus scrofa
18	351	15.6	510	1	ENP1_HUMAN	P49961 homo sapien
19	348.5	15.5	511	1	ENP1_RAT	P97687 rattus norv
20	347.5	15.4	529	1	ENP3_HUMAN	O75355 homo sapien
21	339.5	15.1	495	1	ENP2_HUMAN	Q9y513 homo sapien
22	309.5	13.8	495	1	ENP2_MOUSE	O55026 mus musculu
23	302.5	13.4	495	1	ENP2_RAT	O35795 rattus norv
24	301.5	13.4	494	1	ENP2_CHICK	P79784 gallus gall
25	140	6.2	628	1	NTP1_TOXGO	Q27893 toxoplasma
26	138	6.1	628	1	NTP2_TOXGO	Q27895 toxoplasma
27	131	5.8	592	1	NTP4_TOXGO	P52913 toxoplasma
28	111.5	5.0	634	1	SELB_MOOTH	Q46455 moorella th
29	106.5	4.7	553	1	VG1_SPV4	P11333 spiroplasma
30	104.5	4.6	4427	1	PKSL_BACSU	Q05470 bacillus su
31	101.5	4.5	1151	1	Y245_TREPA	O83273 treponema p
32	99.5	4.4	535	1	DEXB_STRPN	Q54796 streptococc
33	99.5	4.4	548	1	THER_BACST	P06874 bacillus st
34	97	4.3	1220	1	PTC1_BRARE	Q98864 brachydanio
35	96	4.3	711	1	DPP3_YEAST	Q08225 saccharomyc
36	95.5	4.2	2076	1	FAS1_YARLI	P34229 y fatty aci
37	93	4.1	969	1	PAC4_HUMAN	P29122 homo sapien
38	93	4.1	1696	1	ITN2_HUMAN	Q9nzm3 homo sapien
39	92	4.1	290	1	HEM6_CAUCR	Q9aat8 caulobacter
40	91	4.0	417	1	GLYA_BORBU	O51547 borrelia bu
41	91	4.0	477	1	URT1_DESRO	P98119 desmodus ro
42	90	4.0	434	1	YFEW_ECO57	Q8xbj0 escherichia
43	90	4.0	434	1	YFEW_ECOLI	P77619 escherichia
44	89.5	4.0	809	1	RSP5_YEAST	P39940 saccharomyc
45	89.5	4.0	2334	1	WAPA_BACSU	Q07833 bacillus su

ALIGNMENTS

RESULT 1
ENP5_HUMAN
ID ENP5_HUMAN STANDARD; PRT; 428 AA.
AC O75356;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5 precursor
DE (EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like
DE 4) (ER-UDPase).
GN ENTPD5 OR CD39L4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukemia;
RX MEDLINE=98341119; PubMed=9676430;
RA Chadwick B.P., Frischauf A.-M.;
RT "The CD39-like gene family: identification of three new human members
RT (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of
RT the gene family from Drosophila melanogaster.";
RL Genomics 50:357-367(1998).
CC -!- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN
CC GLYCOPROTEINS FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC
CC RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER
CC NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE PYROPHOSPHATE
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC nucleotide + phosphate.
CC -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT LIVER, KIDNEY, PROSTATE,
CC TESTIS AND COLON. MUCH WEAKER EXPRESSION IN OTHER TISSUES.
CC -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF039918; AAC39885.1; -.
DR Genew; HGNC:3367; ENTPD5.
DR MIM; 603162; -.
DR GO; GO:0004002; F:adenosinetriphosphatase activity; TAS.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE_NEG.
KW Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW Endoplasmic reticulum; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 428 ECTONUCLEOSIDE TRIPHOSPHATE

```
FT                              DIPHOSPHOHYDROLASE 5.
FT   CARBOHYD     42     42       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    232    232       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    368    368       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   428 AA;  47517 MW;  830437A155DE4DDD CRC64;

  Query Match             100.0%;  Score 2250;  DB 1;  Length 428;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-173;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 2
ENP5_MOUSE
ID   ENP5_MOUSE     STANDARD;      PRT;   427 AA.
AC   Q9WUZ9; O70214;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 5 precursor
DE   (EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like
DE   4) (ER-UDPase).
GN   ENTPD5 OR CD39L4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=98119025; PubMed=9457681;
RA   Chadwick B.P., Williamson J., Sheer D., Frischauf A.-M.;
RT   "cDNA cloning and chromosomal mapping of a mouse gene with homology to
RT   NTPases.";
RL   Mamm. Genome 9:162-164(1998).
RN   [2]
RP   SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND CHARACTERIZATION.
RC   TISSUE=Liver;
RX   MEDLINE=99298181; PubMed=10369669;
RA   Trombetta E.S., Helenius A.;
RT   "Glycoprotein reglucosylation and nucleotide sugar utilization in the
RT   secretory pathway: identification of a nucleoside diphosphatase in the
RT   endoplasmic reticulum.";
RL   EMBO J. 18:3282-3292(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN
CC       GLYCOPROTEINS FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC
CC       RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER
CC       NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE
CC       PYROPHOSPHATE.
CC   -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC       nucleotide + phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
CC   -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC   -!- TISSUE SPECIFICITY: Ubiquitous.
CC   -!- PTM: GLYCOSYLATED WITH HIGH MANNOSE N-LINKED GLYCANS.
CC   -!- MISCELLANEOUS: OPTIMAL PH IS NEUTRAL.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF006482; AAC05181.1; -.
DR   EMBL; AJ238636; CAB45533.1; -.
DR   EMBL; AK002618; BAB22234.1; -.
DR   MGD; MGI:1321385; Entpd5.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE_NEG.
KW   Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW   Endoplasmic reticulum; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    427       ECTONUCLEOSIDE TRIPHOSPHATE
FT                                DIPHOSPHOHYDROLASE 5.
FT   CARBOHYD     41     41       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    231    231       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    390    390       F -> L (IN REF. 1).
FT   CONFLICT    394    427       DGTLLQLTKKVNNIETGWALGATFHLLQSLGITS -> ERH
FT                                PLTAHKESEQHRDWLGLGGHLSPAPVSGHHQLRPSSTSEAC
FT                                ISEPVFSQEGVDSETFSDLSGKAWPETR (IN REF. 1).
SQ   SEQUENCE   427 AA;  47101 MW;  2F9DA2C342C55577 CRC64;

  Query Match             88.6%;  Score 1994.5;  DB 1;  Length 427;
  Best Local Similarity   88.3%;  Pred. No. 1.1e-152;
  Matches  377;  Conservative   24;  Mismatches   25;  Indels    1;  Gaps    1;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
```

```
Db          1 MATSWGAV-FMLIIACVGSTVFYREQQTWFEGVFLSSMCPINVSAGTFYGIMFDAGSTGT 59

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              |||||||||  |||| ||||:||||||||||||||||||||| ||||||||||||||
Db         60 RIHVYTFVQKTAGQLPFLEGEIFDSVKPGLSAFVDQPKQGAETVQELLEVAKDSIPRSHW 119

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ::|||||||||||||||| ||:||| ||:|||: |||||| ||||||||| |||||||||
Db        120 ERTPVVLKATAGLRLLPEQKAQALLLEVEEIFKNSPFLVPDGSVSIMDGSYEGILAWVTV 179

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              |||||||||  |||||||||||||||||||||||||||||||||||||||:|||||
Db        180 NFLTGQLHGRGQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTFKLYTH 239

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              |||||||||||||||||| :|||||||||||||||||||||||||||||||||||:||
Db        240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGF 299

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              |||||||||||:||||||||:  :|||||||||| || :|||||||:|||||||||||
Db        300 EPCYAEVLRVVQGKLHQPEEVRGSAFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 359

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              |||||| :|:||||||||||:|||||||||||||| |:|||||||||||||||||||||
Db        360 EVCDNLGSFSSGSPFLCMDLTYITALLKDGFGFADGTLLQLTKKVNNIETGWALGATFHL 419

Qy        421 LQSLGIS 427
              |||||||
Db        420 LQSLGIT 426


RESULT 3
ENP5_MESAU
ID   ENP5_MESAU     STANDARD;      PRT;   469 AA.
AC   Q9QYC8; Q9QYC9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 5 precursor
DE   (EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like
DE   4) (ER-UDPase) (Proto-oncogene cph).
GN   ENTPD5 OR CD39L4 OR CPH.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99142925; PubMed=9989819;
RA   Velasco J.A., Avila M.A., Notario V.;
RT   "The product of the cph oncogene is a truncated, nucleotide-binding
RT   protein that enhances cellular survival to stress.";
RL   Oncogene 18:689-701(1999).
CC   -!- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN
CC       GLYCOPROTEINS FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC
CC       RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER
CC       NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE PYROPHOSPHATE
CC       (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC       nucleotide + phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN FETAL CELLS AND MOST ADULT
CC       TISSUES.
CC   -!- DISEASE: CPH ONCOGENE HAS TRANSFORMING CAPACITY AND TUMORIGENIC
CC       POTENTIAL.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF084568; AAF22931.1; ALT_TERM.
DR   EMBL; AF084569; AAF22932.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE_NEG.
KW   Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW   Endoplasmic reticulum; Signal; Proto-oncogene.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    469       ECTONUCLEOSIDE TRIPHOSPHATE
FT                                DIPHOSPHOHYDROLASE 5.
FT   CARBOHYD     42     42       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    232    232       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   469 AA;  52125 MW;  03D8A23E0C73474B CRC64;

  Query Match             85.9%;  Score 1933.5;  DB 1;  Length 469;
  Best Local Similarity   86.7%;  Pred. No. 9.8e-148;
  Matches  364;  Conservative   22;  Mismatches   31;  Indels    3;  Gaps    1;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||| || |||:|:::|  | | :| |||||||:||||||| |||||| ||||||||||||
Db          1 MATPWGAVFFLLMIACAGSTVFYREQQTWFEGVFLSSMCPANVSASTFYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              |||||||||  |||| ||||:||||||||||| |||||||:||  ||||||||||||||
Db         61 RIHVYTFVQKAAGQLPFLEGEIFDSVKPGLSAFADQPKQGAQTVHALLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              |:|||||||||||||| ||:||| ||:|||: |||||| |||||||| |||||||||
Db        121 KRTPVVLKATAGLRLLPEQKAQALLLEVEEIFKMSPFLVPDDSVSIMDGSYEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              |||||||||| |||:|||||||||||||||| |||||||| ||||||||||:|||||
Db        181 NFLTGQLHGHSQETMGTLDLGGASTQITFLPQFSKTLEQTPRDYLTSFEMFNSTFKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              |||||||||||:|||||||::  |||||||||||| || :|||||||:||||||||||
Db        301 EPCYAEVLRVVQGKLHQPEEIRGSSFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              |||||||:|||||||||||||||||||||||||||||| |:|||||||||||||   |  ||
Db        361 EVCDNLESFTSGSPFLCMDLSYITALLKDGFGFADGTLLQLTKKVNNIETGWPGG---HL 417


RESULT 4
ENP6_HUMAN
ID   ENP6_HUMAN     STANDARD;      PRT;   484 AA.
AC   O75354; Q9UJD1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 6 (EC 3.6.1.6)
DE   (NTPDase6) (CD39 antigen-like 2).
GN   ENTPD6 OR CD39L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Keratinocytes;
RX   MEDLINE=98341119; PubMed=9676430;
RA   Chadwick B.P., Frischauf A.-M.;
RT   "The CD39-like gene family: identification of three new human members
RT   (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of
```

```
RT   the gene family from Drosophila melanogaster.";
RL   Genomics 50:357-367(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
CC   -!- FUNCTION: MIGHT SUPPORT GLYCOSYLATION REACTIONS IN THE GOLGI
CC       APPARATUS AND, WHEN RELEASED FROM CELLS, MIGHT CATALYZE THE
CC       HYDROLYSIS OF EXTRACELLULAR NUCLEOTIDES. HYDROLYZES PREFERENTIALLY
CC       NUCLEOSIDE 5'-DIPHOSPHATES, NUCLEOSIDE 5'-TRIPHOSPHATES ARE
CC       HYDROLYZED ONLY TO A MINOR EXTENT, THERE IS NO HYDROLYSIS OF
CC       NUCLEOSIDE 5'-MONOPHOSPHATES. THE ORDER OF ACTIVITY WITH DIFFERENT
CC       SUBSTRATES IS GDP > IDP >> UDP = CDP >> ADP (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC       nucleotide + phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI. BUT ALSO
CC       OCCURS IN A SOLUBLE EXTRACELLULAR FORM (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MOST TISSUES.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF039916; AAC39883.1; -.
DR   EMBL; AL035252; CAB41571.1; -.
DR   Genew; HGNC:3368; ENTPD6.
DR   MIM; 603160; -.
DR   GO; GO:0004002; F:adenosinetriphosphatase activity; TAS.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE_NEG.
KW   Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW   Signal-anchor; Golgi stack.
FT   DOMAIN        1     39       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     40     60       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       61    484       LUMENAL (POTENTIAL).
FT   CARBOHYD    220    220       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    284    284       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    138    138       V -> L (IN REF. 2).
FT   CONFLICT *  202    202       E -> K (IN REF. 2).
SQ   SEQUENCE   484 AA;  53233 MW;  27334B290DB8D64C CRC64;

  Query Match             44.4%;  Score 999;  DB 1;  Length 484;
  Best Local Similarity   52.4%;  Pred. No. 1.2e-72;
  Matches  204;  Conservative   56;  Mismatches  123;  Indels    6;  Gaps    4;

Qy         40 PINVSA---STLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQ 96
              |:  :|      |||||||||||||:||: |  : | : | |  | | :|||||||: |
Db         90 PLGTAADGHEVFYGIMFDAGSTGTRVHVFQFT-RPPRETPTLTHETFKAVKPGLSAYADD 148

Qy         97 PKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSP 156
              :: |: :: ||:|||  ||   || ||:|||||||||||||  ||: || :|||:|: ||
Db        149 VEKSAQGIRELLDVAKQDIPFDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASP 208

Qy        157 FLVPKGSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKT 216
              |||    ||||:|:|||: ||:|:||||| |      :|| ||||| |||| |||: | |
Db        209 FLVGDDCVSIMNGTDEGVSAWITINFLTGSLKTPGGSSVGMLDLGGGSTQIAFLPRVEGT 268

Qy        217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPR 275
              |: :| ||||:  ||| ||||::|||| || :|||| || :|: :   ||    | ||
Db        269 LQASPPGYLTALRMFNRTYKLYSYSYLGLGLMSARLAILGGVEGQPAKDGKELVSPCLSP 328

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDR 335
              :.||    | |: | :      | | | |:: ::|: |||:   |||||||||
Db        329 SFKGEWEHAEVTYRVSGQKAAASLHELCAARVSEVLQNRVHRTEEVKHVDFYAFSYYYDL 388

Qy        336 AVDTDMIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395
              |    :|| |||| | | |||  |: ||  ||     ||| ||||:|:: ||:: |||
Db        389 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTLETQPQSSPFSCMDLTYVSLLLQE-FGFPR 447

Qy        396 STVLQLTKKVNNIETGWALGATFHLLQSL 424
              | ||:||:|::|:|| ||||| || : ||
Db        448 SKVLKLTRKIDNVETSWALGAIFHYIDSL 476


RESULT 5
ENP6_RAT
ID   ENP6_RAT       STANDARD;      PRT;   455 AA.
AC   Q9ER31;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 6 (EC 3.6.1.6)
DE   (NTPDase6) (CD39 antigen-like 2).
GN   ENTPD6 OR CD39L2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND CHARACTERIZATION.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=20498744; PubMed=11042118;
RA   Braun N., Fengler S., Ebeling C., Servos J., Zimmermann H.;
RT   "Sequencing, functional expression and characterization of rat
RT   NTPDase6, a nucleoside diphosphatase and novel member of the ecto-
RT   nucleoside triphosphate diphosphohydrolase family.";
RL   Biochem. J. 351:639-647(2000).
CC   -!- FUNCTION: MIGHT SUPPORT GLYCOSYLATION REACTIONS IN THE GOLGI
CC       APPARATUS AND, WHEN RELEASED FROM CELLS, MIGHT CATALYZE THE
CC       HYDROLYSIS OF EXTRACELLULAR NUCLEOTIDES. HYDROLYZES PREFERENTIALLY
CC       NUCLEOSIDE 5'-DIPHOSPHATES, NUCLEOSIDE 5'-TRIPHOSPHATES ARE
CC       HYDROLYZED ONLY TO A MINOR EXTENT, THERE IS NO HYDROLYSIS OF
CC       NUCLEOSIDE 5'-MONOPHOSPHATES. THE ORDER OF ACTIVITY WITH DIFFERENT
CC       SUBSTRATES IS GDP > IDP >> UDP = CDP >> ADP.
CC   -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC       nucleotide + phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI. BUT ALSO
CC       OCCURS IN A SOLUBLE EXTRACELLULAR FORM.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ277748; CAC16598.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE_NEG.
KW   Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW   Signal-anchor; Golgi stack.
FT   DOMAIN        1     12       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     13     32       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       33    455       LUMENAL (POTENTIAL).
FT   CARBOHYD    192    192       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    256    256       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   455 AA;  49899 MW;  19A22E8BAEF0F77B CRC64;

  Query Match             43.8%;  Score 986;  DB 1;  Length 455;
  Best Local Similarity   54.0%;  Pred. No. 1.2e-71;
  Matches  204;  Conservative   51;  Mismatches  117;  Indels    6;  Gaps    5;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLL 108
              ||||||||||||||||: |  : ||: | |  | | ::||||||: |  :: |: :| ||
Db         74 YGIMFDAGSTGTRIHVFQFA-RPPGETPTLTHETFKALKPGLSAYADDVEKSAQGIQELL 132

Qy        109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMD 168
              |||  ||   || ||:||||||||||||  ||: || :|||:|: |||||    ||||:
Db        133 NVAKQHIPYDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASPFLVGDDCVSIMN 192

Qy        169 GSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQPEKTLEQTPRGYLTSF 228
              |:|||: ||:|||||| |      :|| ||||| ||||||||: | ||: :| |:||:
Db        193 GTDEGVSAWITVNFLTGSLKTPGSSSVGMLDLGGGSTQITFLPRVEGTLQASPPGHLTAL 252

Qy        229 EMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPRWLEAEWIFGGVK 287
              :||| |:|||::|||| || :|||| || :| : . |      | ||      :|    |
Db        253 QMFNRTFKLYSYSYLGLGLMSARLAILGGVEGKPAEDDKELVSPCLSPRFRGKWEHAEVT 312

Qy        288 YQYGGNQEGEVG-FEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEK 346
              |:  | :   || :| | : |  |:| |:|: || |   |||||||| |     :|| ||
Db        313 YRISGQK--AVGLYELCASRVSEVLRNKVHRTEEAQHVDFYAFSYYYDLAASFGLIDAEK 370

Qy        347 GGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVN 406
              || | | |||   |: ||  ||     ||| ||||:||: || : |||    ||:| :|::
Db        371 GGSLVVGDFEIAAKYVCRTLETQPPSSPFACMDLTYISLLLHE-FGFPGDKVLKLARKID 429

Qy        407 NIETGWALGATFHLLQSL 424
              |:|| ||||| || : ||
Db        430 NVETSWALGAIFHYIDSL 447


RESULT 6
NTPA_PEA
ID   NTPA_PEA       STANDARD;      PRT;   455 AA.
AC   P52914;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Nucleoside-triphosphatase (EC 3.6.1.15) (Nucleoside triphosphate
DE   phosphohydrolase) (NTPase) (Apyrase).
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX   NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Alaska; TISSUE=Plumule;
RX   MEDLINE=96197404; PubMed=8616230;
RA   Hsieh H., Tong C.G., Thomas C., Roux S.J.;
RT   "Light-modulated abundance of an mRNA encoding a calmodulin-regulated,
RT   chromatin-associated NTPase in pea.";
RL   Plant Mol. Biol. 30:135-147(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Alaska; TISSUE=Stem;
RA   Shibata K., Abe S., Davies E.;
RT   "Structure of the coding region and mRNA variants of the apyrase from
RT   Pisum sativum.";
RL   Acta Physiol. Plant. 20:3-13(2001).
CC   -!- FUNCTION: MIGHT BE INVOLVED IN RNA TRANSPORT OUT OF NUCLEI.
CC   -!- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + phosphate.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z32743; CAA83655.1; -.
DR   EMBL; AB027614; BAB18900.1; -.
DR   PIR; S65147; S48859.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Nuclear protein.
SQ   SEQUENCE   455 AA;  50072 MW;  50FDF0023ABC4299 CRC64;

  Query Match             21.7%;  Score 489;  DB 1;  Length 455;
  Best Local Similarity   33.7%;  Pred. No. 1e-31;
  Matches  140;  Conservative   62;  Mismatches  160;  Indels   54;  Gaps   14;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEG-EVFDSVKPGLSAFVDQPKQGAETVQGL 107
              | ::|||||||:||||| | | :   | | :| | :: : ||||:: : |:| |:::  |
Db         44 YAVVFDAGSTGSRIHVYHFNQNL-DLLHIGKGVEYYNKITPGLSSYANNPEQAAKSLIPL 102

Qy        108 LEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIF-RKSPFLVPKGSVSI 166
              || |:| :|     ||||  | ||||||||     :: :|  |:::   :| | |   :|||
Db        103 LEQAEDVVPDDLQPKTPVRLGATAGLRLLNGDASEKILQSVRDMLSNRSTFNVQPDAVSI 162

Qy        167 MDGSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQPEKTLEQTPR---- 222
              :||: ||   |||||:  | |     :||| :|||| | |: :    :|| :  |:
Db        163 IDGTQEGSYLWVTVNYALGNLGKKYTKTVGVIDLGGGSVQMAYAVS-KKTAKNAPKVADG 221

Qy        223 --GYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRS--ACLPRWLE 278
                |:    :    | || |||| || :|:|   |              | ||   ||
Db        222 DDPYIKKVVLKGIPYDLYVHSYLHFGREASRAEILKL---------TPRSPNPCLLAGFN 272

Qy        279 AEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSF------------ 326
              : : | :::      |   | |   : : :  ||: |   |  :|
Db        273 GIYTYSGEEFKATAYTSG-ANFNKCKNTIRKAL--KLNYPCPYQNCTFGGIWNGGGGNGQ 329

Qy        327 ---YAFSYYYDRAVDTDMIDYEKGG-ILKVEDFERKAREVCD-NLENFTSGSPFL----- 376
                 :| | ::    || |:|      ||:  | | ||:| |  | |:  |  |||
Db        330 KNLFASSSFFYLPEDTGMVDASTPNFILRPVDIETKAKEACALNFEDAKSTYPFLDKKNV 389

Qy        377 ----CMDLSYITALLKDGFGFADSTVLQLTKKVNN----IETGWALGATFHLLQSL 424
                   |||| |   || ||||      :   |::      :|  | ||     : :|
Db        390 ASYVCMDLIYQYVLLVDGFGLDPLQKITSGKEIEYQDAIVEAAWPLGNAVEAISAL 445


RESULT 7
GDA1_YEAST
ID   GDA1_YEAST     STANDARD;      PRT;   518 AA.
AC   P32621;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Guanosine-diphosphatase (EC 3.6.1.42) (GDPase).
```

```
GN   GDA1 OR YEL042W OR SYGP-ORF16.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=G2-9;
RX   MEDLINE=93308137; PubMed=8391537;
RA   Abeijon C., Yanagisawa K., Mandon E.C., Haeusler A., Moremen K.,
RA   Hirschberg C.B., Robbins P.W.;
RT   "Guanosine diphosphatase is required for protein and sphingolipid
RT   glycosylation in the Golgi lumen of Saccharomyces cerevisiae.";
RL   J. Cell Biol. 122:307-323(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   PubMed=9169868;
RA   Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA   Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA   Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA   Hunicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA   Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA   Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA   Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
RL   Nature 387:78-81(1997).
CC   -!- FUNCTION: AFTER TRANSFER OF SUGARS TO ENDOGENOUS MACROMOLECULAR
CC       ACCEPTORS, THE ENZYME CONVERTS NUCLEOSIDE DIPHOSPHATES TO
CC       NUCLEOSIDE MONOPHOSPHATES WHICH IN TURN EXIT THE GOLGI LUMEN IN
CC       A COUPLED ANTIPORTER REACTION, ALLOWING ENTRY OF ADDITIONAL
CC       NUCLEOTIDE SUGAR FROM THE CYTOSOL.
CC   -!- CATALYTIC ACTIVITY: GDP + H(2)O = GMP + phosphate.
CC   -!- PATHWAY: Glycosylation.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L19560; AAA34656.1; -.
DR   EMBL; U18779; AAB65000.1; -.
DR   PIR; A40732; A40732.
DR   SGD; S0000768; GDA1.
DR   GO; GO:0004382; F:guanosine diphosphatase activity; IDA.
DR   GO; GO:0045134; F:uridine diphosphatase activity; IDA.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Golgi stack; Glycoprotein; Transmembrane; Signal-anchor.
FT   DOMAIN        1      9       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     10     24       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       25    518       LUMENAL (POTENTIAL).
FT   CARBOHYD     41     41       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    335    335       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   518 AA;  56821 MW;  9A61182D5ED22ADF CRC64;

  Query Match             21.5%;  Score 483.5;  DB 1;  Length 518;
  Best Local Similarity   32.8%;  Pred. No. 3.3e-31;
  Matches  142;  Conservative   56;  Mismatches  160;  Indels   75;  Gaps   15;

Qy         49 YGIMFDAGSTGTRIHVYTF-VQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGL 107
              | || ||||||:|:|:| | |   |  | |  | || ::||||:|      | ::  |
Db         93 YVIMIDAGSTGSRVHIYKFDVCTSP---PTLLDEKFDMLEPGLSSFDTDSVGAANSLDPL 149

Qy        108 LEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKS-PFLVPKG-SVS 165
```

```
              |:|| : :|      ||| :|||||||| : |:  :|  |::   |  || | :|  ||
Db        150 LKVAMNYVPIKARSCTPVAVKATAGLRLLGDAKSSKILSAVRDHLEKDYPFPVVEGDGVS 209

Qy        166 IMDGSDEGILAWVTVNFLTGQL--HGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRG 223
              || | :||: ||:| |:| | :  :| :  |    |||| |||| | | |    :
Db        210 IMGGDEEGVFAWITTNYLLGNIGANGPKLPTAAVFDLGGGSTQIVFEPTFPINEKMVDGE 269

Qy        224 YLTSFEMFNSTYKLYTHSYLGFGLKAAR----------LATLGALETEGTDGHTFRSAC 272
              :    :  :  | ||  |:||:|||  |            |      |: : |  |  | |
Db        270 HKFDLKFGDENYTLYQFSHLGYGLKEGRNKVNSVLVENALKDGKILKGDNTKTHQLSSPC 329

Qy        273 LPRWLEA--EWI----------FGGVKYQYGG----------NQEGEVGFEPCYAEVLR 309
              ||  : |  | :           | |      |           |:: :    ||
Db        330 LPPKVNATNEKVTLESKETYTIDFIGPDEPSGAQCRFLTDEILNKDAQCQSPPCSF---- 385

Qy        310 VVRGKLHQPEEV----QRGSFYAFSYYYDRAVDTDM-IDYEKGGILKVEDFERKAREVC- 363
                  :|||  |    :    | |||:|||     | : :       : :    || ||
Db        386 ---NGVHQPSLVRTFKESNDIYIFSYFYDRTRPLGMPLSF------TLNELNDLARIVCK 436

Qy        364 -------------DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIET 410
                          :|:   | | | |:|||:  :||  |:      |:  ||: | |
Db        437 GEETWNSVFSGIAGSLDELESDSHF-CLDLSFQVSLLHTGYDIPLQRELRTGKKIANKEI 495

Qy        411 GWALGATFHLLQS 423
              || |||:  ||::
Db        496 GWCLGASLPLLKA 508
```

```
RESULT 8
APY_SOLTU
ID   APY_SOLTU      STANDARD;      PRT;   454 AA.
AC   P80595; Q43164;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Apyrase precursor (EC 3.6.1.5) (ATP-diphosphatase) (Adenosine
DE   diphosphatase) (ADPase) (ATP-diphosphohydrolase).
GN   RROP1.
OS   Solanum tuberosum (Potato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4113;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 59-160; 236-253 AND 332-345.
RC   TISSUE=Tuber;
RX   MEDLINE=96158985; PubMed=8579614;
RA   Handa M., Guidotti G.;
RT   "Purification and cloning of a soluble ATP-diphosphohydrolase
RT   (apyrase) from potato tubers (Solanum tuberosum).";
RL   Biochem. Biophys. Res. Commun. 218:916-923(1996).
RN   [2]
RP   SEQUENCE OF 42-54; 68-95 AND 236-253.
RC   STRAIN=cv. Desiree;
RX   MEDLINE=96355615; PubMed=8703025;
RA   Vasconcelos E.G., Ferreira S.T., de Carvalho T.M.U., de Souza W.,
RA   Kettlun A.M., Mancilla M., Valenzuela M.A., Verjovski-Almeida S.;
RT   "Partial purification and immunohistochemical localization of ATP
RT   diphosphohydrolase from Schistosoma mansoni. Immunological cross-
RT   reactivities with potato apyrase and Toxoplasma gondii nucleoside
RT   triphosphate hydrolase.";
RL   J. Biol. Chem. 271:22139-22145(1996).
CC   -!- FUNCTION: CATALYZES THE HYDROLYSIS OF PHOSPHOANHYDRIDE BONDS OF
CC       NUCLEOSIDE TRI- AND DI-PHOSPHATES.
CC   -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC   -!- COFACTOR: CALCIUM.
CC   -!- SUBCELLULAR LOCATION: Membrane-associated (Probable).
CC   -!- PTM: THE N-TERMINUS IS BLOCKED.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U58597; AAB02720.1; -.
DR   PIR; JC4616; JC4616.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Transmembrane; Calcium; Signal.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    454       APYRASE.
FT   TRANSMEM    426    446       POTENTIAL.
FT   CARBOHYD    151    151       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    262    262       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   454 AA;  50041 MW;  9D9EFE431DA2F52F CRC64;

  Query Match             19.8%;  Score 446.5;  DB 1;  Length 454;
  Best Local Similarity   28.8%;  Pred. No. 2.6e-28;
  Matches  131;  Conservative   78;  Mismatches  177;  Indels   69;  Gaps   16;

Qy         25 NQQTWFEGIFLS-------SMCPINVSA-----------STLYGIMFDAGSTGTRIHVY 65
              || : |  | |:       |:   ||:|           |  | ::|||||||:|:||:
Db          3 NQNSHFIFIILAIFLVLPLSLLSKNVNAQIPLRRHLLSHESEHYAVIFDAGSTGSRVHVF 62

Qy         66 TFVQKMPGQLPILEG-EVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTP 124
                | :|: | |||     | | : :||||:: : ||  | ::: ||: |:  :|:    :||
Db         63 RFDEKL-GLLPIGNNIEYFMATEPGLSSYAEDPKAAANSLEPLLDGAEGVVPQELQSETP 121

Qy        125 VVLKATAGLRLLPEHKAKALLFEVKEIFR-KSPFLVPKGSVSIMDGSDEGILAWVTVNFL 183
              : | |||||||:|     |: :|  |: : : :| |       |:|:||: ||   |  :|:|
Db        122 LELGATAGLRMLKGDAAEKILQAVRNLVKNQSTFHSKDQWVTILDGTQEGSYMWAAINYL 181

Qy        184 TGQLHGHRQETVGTLDLGGASTQITFL---PQFEKTLE-QTPRGYLTSFEMFNSTYKLYT 239
                | |    : |  |:|||| | |: :     || |  : :    |:    : :  | ||
Db        182 LGNLGKDYKSTTATIDLGGGSVQMAYAISNEQFAKAPQNEDGEPYVQQKHLMSKDYNLYV 241

Qy        240 HSYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVG 299
              |||| :|  | |     |- |·  :  |      :  : :||| |:     ::|
Db        242 HSYLNYGQLAGRAEIFKASRNES-------NPCALEGCDGYYSYGGVDYKVKAPKKGS-- 292

Qy        300 FEPCYAEVLRVVRG--KLHQPEEVQRGSF--------------YAFSYYYDRAVDTDMI 342
                  :    |: |   |::    ::  :|              :| |::|| ::
Db        293 ---SWKRCRRLTRHALKINAKCNIEECTFNGVWNGGGGDGQKNIHASSFFYDIGAQVGIV 349

Qy        343 DYE-KGGILKVEDFERKAREVCD-NLENFTS--------GSPFLCMDLSYITALLKDGFG 392
              | :    : |   :   |:  |  |: :  |          |:|||| |   || ||||
Db        350 DTKFPSALAKPIQYLNAAKVACQTNVADIKSIFPKTQDRNIPYLCMDLIYEYTLLVDGFG 409

Qy        393 F---ADSTVLQLTKKVNN-IETGWALGATFHLLQS 423
                : ||:   :  |  :    | ||    |: |
Db        410 LNPHKEITVIHDVQYKNYLVGAAWPLGCAIDLVSS 444


RESULT 9
YY4E_CAEEL
ID   YY4E_CAEEL     STANDARD;      PRT;   485 AA.
AC   Q18411;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 54.3 kDa protein C33H5.14 in chromosome IV.
GN   C33H5.14.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Bradshaw H., Stellyes L.;
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U41007; AAA82272.1; -.
DR   PIR; T34147; T34147.
DR   WormPep; C33H5.14; CE04157.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hypothetical protein; Transmembrane; Hydrolase.
FT   TRANSMEM    439    459       POTENTIAL.
SQ   SEQUENCE   485 AA;  54309 MW;  DE64D1ADC20F581E CRC64;

  Query Match             17.2%;  Score 387.5;  DB 1;  Length 485;
  Best Local Similarity   27.7%;  Pred. No. 1.6e-23;
  Matches  120;  Conservative   84;  Mismatches  152;  Indels   77;  Gaps   18;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDS------VKPGLSAFVDQPKQGAE 102
              ||:: ||||:|||: ||| :: : | |  ::  : :|       | ||||:| |:|:|  |
Db         25 YGVICDAGSSGTRLFVYT-LKPLSGGLTNIDTLIHESEPVVKKVTPGLSSFGDKPEQVVE 83

Qy        103 TVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFL-VPK 161
                :  ||  |:: ||      :| ::: ||||:|||||  :  |::  ::   :    | |
Db         84 YLTPLLRFAEEHIPYEQLGETDLLIFATAGMRLLPEAQKDAIIKNLQNGLKSVTALRVSD 143

Qy        162 GSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEK------ 215
              :: |:||: ||| :|: ||:: |:          || :|:|||| || |    ||
Db        144 SNIRIIDGAWEGIYSWIAVNYILGRFDKENDSKVGMIDMGGASVQIAFEIANEKESYNGG 203

Qy        216 TLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRSACLPR 275
              : :   | : : | :   ||:|: ::||:|          :|  |       :| ||
Db        204 NVYEINLGSIETNEDYK--YKIYSTTFLGYGANEGLKKYENSLVKSGNS----NDSCSPR 257

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPE---------------- 319
              |   : |    ::  |  ||  :: | |:|  :: |   ||
Db        258 GLNR--LIG----EFTVNGTGE--WDVCLAQVSSLI-GDKAQPSCPNPTCFLRNVIAPSV 308

Qy        320 EVQRGSFYAFS-YYYDRAVDTDMIDYEKGGILKVEDFERKAREVC----DNLENFTSGSP 374
              :      | || |:|  :      ::  ||    : |  : |:|    :::::    :
Db        309 NLSTVQLYGFSEYWYTTS------NFGSGGEYHYQKFTDEVRKYCQKDWNDIQDGFKRNE 362

Qy        375 F----------LCMDLSYITALLKDGFGFADST--VLQLTKKVNNIETGWALGA------ 416
              |          |   :::|::| |||   | |  : |   |:   |  |||||
Db        363 FPNADIERLGTNCFKAAWVTSVLHDGFN-VDKTKHLFQSVLKIAGEEMQWALGAMLYHSK 421

Qy        417 --TFHLLQSLGIS 427
                |:||: | ::
Db        422 DLKFNLLEQLEVA 434


RESULT 10
YBU4_CAEEL
ID   YBU4_CAEEL     STANDARD;      PRT;   552 AA.
AC   Q21815;
DT   01-NOV-1997 (Rel. 35, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein R07E4.4 in chromosome X.
GN   R07E4.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
```

```
RA   Miller N.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   REVISIONS.
RA   Waterston R.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U39652; AAA80403.2; -.
DR   WormPep; R07E4.4; CE28748.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hypothetical protein; Transmembrane; Hydrolase.
FT   TRANSMEM      7     27       POTENTIAL.
FT   TRANSMEM    490    510       POTENTIAL.
SQ   SEQUENCE   552 AA;  62510 MW;  E5DC32C858AE4D94 CRC64;

  Query Match           17.1%;  Score 385;  DB 1;  Length 552;
  Best Local Similarity  28.6%;  Pred. No. 3e-23;
  Matches  126;  Conservative   69;  Mismatches  160;  Indels   86;  Gaps   20;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDS------VKPGLSAFVDQPKQGAE 102
              ||:: ||||||||: || ::     :|  :|  ::|:      : |||| |  :| | ||
Db         44 YGVICDAGSTGTRLFVYNWISTSDSELIQIEPVIYDNKPVMKKISPGLSTFGTKPAQAAE 103

Qy        103 TVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFL---- 158
               :: |:|:|:  ||      ||| : ||||:||:|: : :|:|   | :  | | :
Db        104 YLRPLMELAERHIPEEKRPYTPVFIFATAGMRLIPDEQKEAVL---KNLRNKLPKITSMQ 160

Qy        159 VPKGSVSIMDGSDEGILAWVTVNFLTGQLH-----------GH-RQETVGTLDLGGAST 205
              | |  : |::|  ||| :|: ||:  |: :             | ||:||| :|:||||
Db        161 VLKEHIRIIEGKWEGIYSWIAVNYALGKFNKTATLDFPGTSPAHARQKTVGMIDMGGASA 220

Qy        206 QITF-LPQFEK----TLEQTPRGYLTSFEMFNSTYKLYTHSYLGF----GLKAARLATLG 256
              || | ||  :      :|    |      :|   |||:  ::||:    |::      |
Db        221 QIAFELPDTDSFSSINVENINLGCREDDSLFK--YKLFVTTFLGYGVNEGIRKYEHMLLS 278

Qy        257 ALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGN--QEG----------------E 297
              |: :  :|   :  |:|  |        |  : | |  : |                  |
Db        279 KLKDQ--NGTVIQDDCMPLNLHKT-----VTLENGENFVRRGTGNWNTCSNEVKKLLNPE 331

Qy        298 VGFEPCYAEVLRVVRGKLHQPE-EVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFE 356
              | | ||   :   | :  |   :      | || |:     |      |     ||    |:
Db        332 SSSEVCKAEAAKCYFGAVPAPSIPLSNIEMYGFSEYWYSTHDV----LGLGGQYDAENIA 387

Qy        357 RKAREVCD------NLENFTSGSP--------FLCMDLSYITALLKDGFGFADST--VLQ 400
              :| :: |        |:     |          |   ::||::| |||   | |    |
Db        388 KKTQQYCSKRWSTIQAESKKQLYPRADEERLRTQCFKSAWITSVLHDGFS-VDKTHNKFQ 446

Qy        401 LTKKVNNIETGWALGA-TFHL 420
                :   |  |||||  :|:
Db        447 SVSTIAGQEVQWALGAMIYHM 467


RESULT 11
ENP1_CHICK
ID   ENP1_CHICK     STANDARD;      PRT;   493 AA.
AC   O93295;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE   (NTPDase1) (Ecto-ATP diphosphohydrolase) (ATPDase) (Lymphoid cell
DE   activation antigen) (Ecto-apyrase) (CD39 antigen).
GN   ENTPD1 OR CD39.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-21 AND 150-156.
RC   TISSUE=Oviduct;
RX   MEDLINE=98298108; PubMed=9632655;
RA   Nagy A.K., Knowles A.F., Nagami G.T.;
RT   "Molecular cloning of the chicken oviduct ecto-ATP-
RT   diphosphohydrolase.";
RL   J. Biol. Chem. 273:16043-16049(1998).
RN   [2]
RP   SEQUENCE OF 1-17.
RC   TISSUE=Stomach;
RX   MEDLINE=97442428; PubMed=9295305;
RA   Lewis-Carl S., Kirley T.L.;
RT   "Immunolocalization of the ecto-ATPase and ecto-apyrase in chicken
RT   gizzard and stomach. Purification and N-terminal sequence of the
RT   stomach ecto-apyrase.";
RL   J. Biol. Chem. 272:23645-23652(1997).
CC   -!- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC       nucleotides to regulate purinergic neurotransmission. Could also
CC       be implicated in the prevention of platelet aggregation.
CC       Hydrolyzes ATP and ADP equally well (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- PTM: N-GLYCOSYLATED.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF041355; AAC26491.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT   DOMAIN        1      7       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      8     28       POTENTIAL.
FT   DOMAIN       29    463       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    464    486       POTENTIAL.
FT   DOMAIN       487    493       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      65     65       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      79     79       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     133    133       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     223    223       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     267    267       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     324    324       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     330    330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     361    361       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     372    372       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     382    382       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT      16     16       C -> W (IN REF. 2).
FT   CONFLICT      21     21       I -> G (IN REF. 1; AA SEQUENCE).
SQ   SEQUENCE   493 AA;  54034 MW;  F14FF4C3AA2F3603 CRC64;

  Query Match           17.0%;  Score 383;  DB 1;  Length 493;
  Best Local Similarity  27.3%;  Pred. No. 3.7e-23;
  Matches  131;  Conservative   66;  Mismatches  170;  Indels  112;  Gaps   18;

Qy         11 MLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYTF-VQ 69
```

```
Db         11 LLTATCVFSIIALILSAVDVKDVFL--------PPGTKYGLVFDAGSTHTALYVYQWPAD 62

Qy         70 KMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKA 129

Db         63 KENGTGIVSQVESCTVNGSGISSYADDPAGAGASLKPCLDKAMAVIPVEQQWQTPTYLGA 122

Qy        130 TAGLRLLPEH---KAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTVNFLTGQ 186

Db        123 TAGMRLLREQNSTKAEQVFAEVSKAIREFP--VDFRGAQILTGNEEGSFGWITVNYLLET 180

Qy        187 L----------HGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYK 236

Db        181 LIKFSFAGKWEHPQNTEVLGALDLGGASTQITFQPGV--TIEDKNTSVL--FRLYGTNYS 236

Qy        237 LYTHSYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQE- 295

Db        237 LYTHSYLCYGQIQASKRLMAALHQDGSYVQNISHPCYPK----------------GYRRI 280

Qy        296 ---GEVGFEPCY--------AEVLRVVR--------------------------GKLHQ 317

Db        281 ITIAEIYDSPCVPTPSMLSPAQILTVTGTGNPAACPTAILKLFNLTCGANRTCGFDGVYQ 340

Qy        318 PEEVQRGSFYAFS-YYYDRA------------VDTDMIDY-EKGGILKVEDFERKAREVC 363

Db        341 PP--VRGQFFAFAGFYYTFSFLNLTGQQSLSHVNATVWDFCNKNWSELVETFPQNK---- 394

Qy        364 DNLENFTSGGSPFLCMDLSYITALLKDGFGFADST--VLQLTKKVNNIETGWALGATFHL 420

Db        395 EHLHTY-------CVVGLYILTLLVDGYKFDEHTWSNIHFSQKAGNADIGWTLGFMLNL 446


RESULT 12
YND1_YEAST
ID   YND1_YEAST     STANDARD;      PRT;   630 AA.
AC   P40009;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Golgi apyrase (EC 3.6.1.5) (ATP-diphosphatase) (Adenosine
DE   diphosphatase) (ADPase) (ATP-diphosphohydrolase) (Golgi nucleoside
DE   diphosphatase).
GN   YND1 OR YER005W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   STRAIN=S288c;
RX   MEDLINE=99340091; PubMed=10409709;
RA   Gao X.D., Kaigorodov V., Jigami Y.;
RT   "YND1, a homologue of GDA1, encodes membrane-bound apyrase required
RT   for Golgi N- and O-glycosylation in Saccharomyces cerevisiae.";
RL   J. Biol. Chem. 274:21450-21456(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   PubMed=9169868;
RA   Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA   Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA   Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA   Hunicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA   Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA   Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA   Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
RL   Nature 387:78-81(1997).
CC   -!- FUNCTION: CATALYZES THE HYDROLYSIS OF PHOSPHOANHYDRIDE BONDS OF
CC       NUCLEOSIDE TRI-AND DI-PHOSPHATES. HAS EQUAL HIGH ACTIVITY TOWARD
CC       ADP/ATP, GDP/GTP, AND UDP/UTP AND APPROXIMATELY 50% LESS TOWARD
CC       CDP/CTP AND THIAMINE PYROPHOSPHATE. HAS NO ACTIVITY TOWARD GMP.
CC       REQUIRED FOR GOLGI GLYCOSYLATION AND CELL WALL INTEGRITY.
CC   -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC   -!- PATHWAY: Glycosylation.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE BOUND; GOLGI.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF203695; AAF17573.1; -.
DR   EMBL; U18778; AAB64538.1; -.
DR   PIR; S50463; S50463.
DR   SGD; S0000807; YND1.
DR   GO; GO:0006486; P:protein amino acid glycosylation; IMP.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Transmembrane; Golgi stack.
FT   DOMAIN        1    500       LUMENAL (POTENTIAL).
FT   TRANSMEM    501    517       POTENTIAL.
FT   DOMAIN      518    630       CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   630 AA;  71851 MW;  02F8D24A78212544 CRC64;

  Query Match             16.5%;  Score 372;  DB 1;  Length 630;
  Best Local Similarity   26.2%;  Pred. No. 3.9e-22;
  Matches  117;  Conservative   73;  Mismatches  165;  Indels   92;  Gaps   17;

Qy         49 YGIMFDAGSTGTRIHVYTF------------VQKMPGQLPILEGE---VFDSVKPGLSA 92

Db         10 FGIVIDAGSSGSRIHVFKWQDTESLLHATNQDSQSILQSVPHIHQEKDWTF-KLNPGLSS 68

Qy         93 FVDQPKQGAET-VQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEV-KE 150

Db         69 FEKKPQDAYKSHIKPLLDFAKNIIPESHWSSCPVFIQATAGMRLLPQDIQSSILDGLCQG 128

Qy        151 IFRKSPFLVP--KGSVSIMDGSDEGILAWVTVNFLTGQLHGHRQE-----TVGTLDLGGA 203

Db        129 LKHPAEFLVEDCSAQIQVIDGETEGLYGWLGLNYLYGHFNDYNPEVSDHFTFGFMDMGGA 188

Qy        204 STQITFLPQFEKTLE----------QTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARL 252

Db        189 STQIAFAPHDSGEIARHRDDIATIFLRSVNGDLQKWDVFVST-------WLGFGANQARR 241

Qy        253 ATLGAL----------ETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGFE 301

Db        242 RYLAQLINTLPENTNDYENDDFSTRNLNDPCMPRGSSTDFEFKDTIFHIAGSGNYEQCTK 301

Qy        302 PCYAEVLR--------VVRGKLHQPE-EVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKV 352

Db        302 SIYPLLLKNMPCDDEPCLFNGVHAPRIDFANDKFIGTSEYWYTANDV----FKLGGEYNF 357

Qy        353 EDFERKAREVCDN-----LENFTSG------SPFL---CMDLSYITALLKDGFGFA---- 394

Db        358 DKFSKSLREFCNSNWTQILANSDKGVYNSIPENFLKDACFKGNWVLNILHEGFDMPRIDV 417

Qy        395 ------DSTVLQLTKKVNNIETGWALG 415

Db        418 DAENVNDRPLFQSVEKVEERELSWTLG 444


RESULT 13
ENP1_MOUSE
ID   ENP1_MOUSE     STANDARD;      PRT;   510 AA.
AC   P55772;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE   (NTPDase1) (Ecto-ATP diphosphohydrolase) (ATPDase) (Lymphoid cell
```

```
DE   activation antigen) (Ecto-apyrase) (CD39 antigen).
GN   ENTPD1 OR CD39.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95015846; PubMed=7930580;
RA   Maliszewski C.R., Delespesse G.J.T., Schoenborn M.A., Armitage R.J.,
RA   Fanslow W.C., Nakajima T., Baker E., Sutherland G.R., Poindexter K.,
RA   Birks C., Alpert A., Friend D., Gimpel S.D., Gayle R.B. III;
RT   "The CD39 lymphoid cell activation antigen. Molecular cloning and
RT   structural characterization.";
RL   J. Immunol. 153:3574-3583(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=98399871; PubMed=9730622;
RA   Schoenborn M.A., Jenkins N.A., Copeland N.G., Gilbert D.J.,
RA   Gayle R.B. III, Maliszewski C.R.;
RT   "Gene structure and chromosome location of mouse Cd39 coding for an
RT   ecto-apyrase.";
RL   Cytogenet. Cell Genet. 81:287-289(1998).
CC   -!- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC       nucleotides to regulate purinergic neurotransmission. Could also
CC       be implicated in the prevention of platelet aggregation.
CC       Hydrolyzes ATP and ADP equally well.
CC   -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC   -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF037366; AAB92259.1; -.
DR   EMBL; AF041818; AAC83203.1; -.
DR   EMBL; AF041812; AAC83203.1; JOINED.
DR   EMBL; AF041813; AAC83203.1; JOINED.
DR   EMBL; AF041814; AAC83203.1; JOINED.
DR   EMBL; AF041815; AAC83203.1; JOINED.
DR   EMBL; AF041816; AAC83203.1; JOINED.
DR   EMBL; AF041817; AAC83203.1; JOINED.
DR   MGD; MGI:102805; Entpd1.
DR   GO; GO:0005605; C:basal lamina; IDA.
DR   GO; GO:0004050; F:apyrase activity; IDA.
DR   GO; GO:0006200; P:ATP catabolism; IDA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IDA.
DR   GO; GO:0030168; P:platelet activation; IDA.
DR   GO; GO:0009181; P:purine ribonucleoside diphosphate catabolism; IDA.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT   DOMAIN        1     16       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     17     37       POTENTIAL.
FT   DOMAIN       38    478       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    479    499       POTENTIAL.
FT   DOMAIN      500    510       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     73     73       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    291    291       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    428    428       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    457    457       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   510 AA;  57205 MW;  8E6A6113D2E13930 CRC64;
```

```
  Query Match             16.4%;  Score 370;  DB 1;  Length 510;
  Best Local Similarity   27.1%;  Pred. No. 4.2e-22;
  Matches  121;  Conservative   71;  Mismatches  173;  Indels   82;  Gaps   19;

Qy         32 GIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVK-PGL 90
              |:  :   | ||     |||: ||||: | :::| :  :      :::       || ||:
Db         36 GLTQNKPLPENVK----YGIVLDAGSSHTNLYIYKWPAEKENDTGVVQQLEECQVKGPGI 91

Qy         91 SAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLL---PEHKAKALLFE 147
              | :  :  :    :   :|:: : || |   :||| | ||||:|||    |  |  :|
Db         92 SKYAQKTDEIGAYLAECMELSTELIPTSKHHQTPVYLGATAGMRLLRMESEQSADEVLAA 151

Qy        148 VKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTVNFLTGQL----------HGHRQETVG 196
              |    :  ||         |: | :||   |:|:|:| |:             :||| |
Db        152 VSTSLKSYPF--DFQGAKIITGQEEGAYGWITINYLLGRFTQEQSWLSLISDSQKQETFG 209

Qy        197 TLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLG 256
              ||||||||||||:||   |:| :|   |  | ::   | :||||:| :|   |    |
Db        210 ALDLGGASTQITFVPQ-NSTIE-SPENSL-QFRLYGEDYTVYTHSFLCYGKDQALWQKL- 265

Qy        257 ALETEGTDGHTFRSACL-----------------PRWLEAEWIFGGVKYQYGGNQEGEV 298
              | : : |   :  |                     :  | :  |   :| |:
Db        266 AKDIQVSSGGVLKDPCFNPGYEKVVNVSELYGTPCTKRFEKKLPFDQFRIQGTGD----- 320

Qy        299 GFEPCYAEVLRVVRGKLHQPEE----------VQRGSFYAFSYYYDRAVDTDMIDY---- 344
              :| |:  :| :     | |              ||| ||| :|       ::|:
Db        321 -YEQCHQSILELFNNS-HCPYSQCAFNGVFLPPLHGSFGAFSAFY------FVMDFFKKV 372

Qy        345 EKGGILKVEDFERKAREVCD-NLENFTSGSPFL--------CMDLSYITALLKDGFGFAD 395
              |  :: |      :  | : |   :  | :          |   :|| :||: |: | |
Db        373 AKNSVISQEKMTEITKNFCSKSWEETKTSYPSVKEKYLSEYCFSGAYILSLLQ-GYNFTD 431

Qy        396 STVLQL--TKKVNNIETGWALGATFHL 420
              |:  |:    |: :   || ||   :|
Db        432 SSWEQIHFMGKIKDSNAGWTLGYMLNL 458
```

```
RESULT 14
ENP1_BOVIN
ID   ENP1_BOVIN     STANDARD;      PRT;   513 AA.
AC   O18956;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE   (NTPDase1) (Ecto-ATP diphosphohydrolase) (ATPDase) (Lymphoid cell
DE   activation antigen) (CD39 antigen) (Ecto-apyrase).
GN   ENTPD1 OR CD39.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Aortic endothelium;
RA   Chang A.S., Garcia R.L., Chang S.M., Schilling W.P.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 97-103; 123-133; 136-140; 145-168 AND 459-471.
RC   TISSUE=Aorta;
RX   MEDLINE=97115858; PubMed=8955160;
RA   Kaczmarek E., Koziak K., Sevigny J., Siegel J.B., Anrather J.,
RA   Beaudoin A.R., Bach F.H., Robson S.C.;
RT   "Identification and characterization of CD39/vascular ATP
RT   diphosphohydrolase.";
RL   J. Biol. Chem. 271:33116-33122(1996).
CC   -!- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC       nucleotides to regulate purinergic neurotransmission. Could also
CC       be implicated in the prevention of platelet aggregation.
CC       Hydrolyzes ATP and ADP equally well.
CC   -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
```

```
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
CC   -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF005940; AAB62382.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT   DOMAIN        1     16       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     17     37       POTENTIAL.
FT   DOMAIN       38    481       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    482    502       POTENTIAL.
FT   DOMAIN      503    513       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     73     73       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    227    227       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    245    245       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    307    307       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    336    336       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    373    373       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    460    460       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     97     97       K -> N (IN REF. 2).
FT   CONFLICT    101    103       INV -> CGF (IN REF. 2).
FT   CONFLICT    464    464       K -> V (IN REF. 2).
SQ   SEQUENCE   513 AA;  58113 MW;  20FE98F27B6D2F96 CRC64;

  Query Match           16.2%;  Score 365;  DB 1;  Length 513;
  Best Local Similarity  26.5%;  Pred. No. 1.1e-21;
  Matches  129;  Conservative   70;  Mismatches  180;  Indels  108;  Gaps   19;

Qy          8 VFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYTF 67
              :  :|  ||: : ::         |:  :   | ||     :||: ||||: | :::| :
Db         17 ILSILGFSCIIAVIA-----LLALGLTQNKALPENVK----FGIVLDAGSSHTSLYIYRW 67

Qy         68 VQKMPGQLPILEGEVFDSVK-PGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTPVV 126
                :      ::      :|| ||:| |  :  :    :  :| |:  ||    :|||
Db         68 PAEKENDTGVVTQIEESNVKGPGISGFAKKVNEINVYLTACMERAQKVIPSIQHMETPVY 127

Qy        127 LKATAGLRLL-PEHK--AKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTVNFL 183
              | ||||:|||  |:|  |  :|  |     : ||          |: | :||   |:|||:|
Db        128 LGATAGMRLLRMENKQMADKILAAVASSISEYPF--DFQGARIISGQEEGAYGWITVNYL 185

Qy        184 TGQL-----------HGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMF 231
              |:                ||| | |||||||||||:|| | |  ::|    |  | ::
Db        186 LGKFTQKLSWFNLKPSKDDTQETYGALDLGGASTQITFVPQNETT--ESPNNNL-YFRLY 242

Qy        232 NSTYKLYTHSYLGFGLKAARLATLGALETEGTDG------------------------ 265
                | :||||:| :|   | |  | ||  :||:|
Db        243 GKNYSVYTHSFLCYGKDQALLQKL-ALGLQGTNGIIHEPCFHSRYMRKIKMSVLNEGFCT 301

Qy        266 --HTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRV-----------V 311
                |   |:  |          :: :  ||      |: |  ::::
Db        302 KRHELNSSFYP--------LVDIEIRGAGN------FQRCRQSIIQLFNTSYCPYSSCSF 347

Qy        312 RGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAREVC-------- 363
                |    |   | |:| || |        : ::      : ||    |  | |
Db        348 NGVFLPPLHGQFGAFSAFYYV------MEFLNLTSEESVSVEQLTEKLEEFCAQRWEEVQ 401

Qy        364 ----DNLENFTSGSPFLCMDLSYITALLKDGFGF-ADS-TVLQLTKKVNNIETGWALGAT 417
                  :  | : |     |  :||  || :|: | |:|   :    || : : || ||
Db        402 KNFGEVKEKYLSE---YCFSGTYILVLLLNGYHFTAESWKNIHFMNKVRSTDVGWTLGYM 458

Qy        418 FHLLQSL 424
                :|  :
Db        459 LNLTNKI 465


RESULT 15
ENP4_MOUSE
ID   ENP4_MOUSE     STANDARD;      PRT;   613 AA.
AC   Q9DBT4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 4 (EC 3.6.1.6)
DE   (NTPDase4) (Uridine-diphosphatase) (UDPase) (Lysosomal apyrase-like
DE   protein of 70 kDa).
GN   LYSAL1 OR ENTPD4 OR LALP70.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=C57BL/6J; TISSUE=Lung;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE OF 5-613 FROM N.A. (ISOFORM 2).
RC   TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX   MEDLINE=20317099; PubMed=10858452;
RA   Biederbick A., Kosan C., Kunz J., Elsaesser H.-P.;
RT   "First apyrase splice variants have different enzymatic properties.";
RL   J. Biol. Chem. 275:19018-19024(2000).
CC   -!- FUNCTION: HYDROLYZES PREFERENTIALLY NUCLEOSIDE 5'-DIPHOSPHATES,
```

```
CC       NUCLEOSIDE 5'-TRIPHOSPHATES ARE HYDROLYZED ONLY TO A MINOR EXTENT
CC       (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC       nucleotide + phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC       LOCALIZES IN THE GOLGI AND AUTOPHAGIC VACUOLES/LYSOSOMES (By
CC       similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1; Synonyms=LALP70;
CC         IsoId=Q9DBT4-1; Sequence=Displayed;
CC       Name=2; Synonyms=LALP70V;
CC         IsoId=Q9DBT4-2; Sequence=VSP_003615;
CC   -!- TISSUE SPECIFICITY: Ubiquitous.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AK004761; BAB23542.1; -.
DR   EMBL; BC006924; AAH06924.1; -.
DR   MGD; MGI:1914714; Lysal1.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE_NEG.
KW   Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW   Alternative splicing; Golgi stack; Lysosome.
FT   DOMAIN        1     33       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     34     54       POTENTIAL.
FT   DOMAIN       55    559       LUMENAL (POTENTIAL).
FT   TRANSMEM    560    580       POTENTIAL.
FT   DOMAIN      581    613       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD    404    404       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    407    407       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    287    294       Missing (in isoform 2).
FT                                /FTId=VSP_003615.
SQ   SEQUENCE   613 AA;  69745 MW;  DE528F512ABEF52F CRC64;

  Query Match             16.0%;  Score 360.5;  DB 1;  Length 613;
  Best Local Similarity   24.4%;  Pred. No. 3.2e-21;
  Matches  127;  Conservative   77;  Mismatches  185;  Indels  131;  Gaps   21;

Qy         11 MLVVSCVCSAVS--------HRNQQTWFE-----GIFLSSMCPINV----SASTLYGIMF 53
              ::|:| : :|||         |::  |         :|: :  :      : |  |||:
Db         34 IVVISILAAAVSLLYFSVVIIRSKYGWLSKDKKFQRYLARVTDVEATDTNNPSVNYGIVV 93

Qy         54 DAGSTGTRIHVYTFVQKMPGQLPILE---------GEVFDSVKPGLSAFVDQPKQGAETV 104
              | ||:|:|| || : :       :|:           |   :|||:| |   |:: :: :
Db         94 DCGSSGSRIFVYCWPRHNGNPHDLLDIRQMRDKNRKPVVMKIKPGISEFATSPEKVSDYI 153

Qy        105 QGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSP----FLVP 160
                ||  | : :||:  |:||: :  |||:|:||| : ||:|   :::    |    ||
Db        154 SPLLSFAAEHVPRAKHKETPLYILCTAGMRVLPESQQKAIL---EDLLTDIPVHYDFLFS 210

Qy        161 KGSVSIMDGSDEGILAWVTVNFLTGQLHGH----------------------RQETVGT 197
                 :: |  ||: ||: :||: |:   |                        |: | |
Db        211 DSHAEVISGKQEGVYAWIGINFVLGRFE-HIEEDDEAVVEVNIPGSESSEAIVRKRTAGV 269

Qy        198 LDLGGASTQITF-LPQ----FEKTLEQTPRGYLTSFEM------FNSTYKLYTHSYLGFG 246
              ||:|| |||| : :||          |:  :  |  | :          |::|  ::||||
Db        270 LDMGGVSTQIAYEVPQTVSFASSQQEEVAKNLLAEFNLGCDVHQTEHVYRVYVATFLGFG 329

Qy        247 LKAARLA------------TLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGN 293
              |||          |  :|  |         |||  :: |       |  |
Db        330 GNAARQRYEDRLFASTVQKNRLLGKQTGLTPDAPLLDPCLPLDIKDE-------IQQNGQ 382

Qy        294 Q---EGEVGFEPCYAEVLRVVRGKLHQ----------PEEVQRGSFYAFSYYYDRAVDT 339
```

```
                 :|   |: |  | |:    | ::          |   |   || || :|    |
Db        383 TLYLQGTGDFDLC-RETLQPFMNKTNETQTSLNGVYQPPIHFQNSEFYGFSEFYYCTEDV 441

Qy        340 DMIDYEKGGILKVEDFERKAREVCDN-----LENFTSG----------SPFLCMDLSYIT 384
                  ||       | : |:: |         | |  |              : |   :::
Db        442 ----LRMGGDYNAARFTQAAKDYCATKWSILRERFDRGLYASHADLHRLKYQCFKSAWMF 497

Qy        385 ALLKDGFGFADS-----TVLQLTKKVNNIETGWALGATFH 419
                  :  || |  :     | ||:  |    |  | |||  :
Db        498 EVFHKGFSFPVTYKNLKTALQVYDK----EVQWTLGAILY 533
```

Search completed: January 29, 2004, 10:04:09
Job time : 11.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 10:05:35 ; Search time 29.5 Seconds
(without alignments)
3015.210 Million cell updates/sec

Title: US-10-091-085-7
Perfect score: 2252
Sequence: 1 MATSWGTVFFMLVVSCVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2252	100.0	428	12	US-10-286-926-7	Sequence 7, Appli
2	2252	100.0	428	14	US-10-091-085-7	Sequence 7, Appli
3	2252	100.0	428	14	US-10-092-063-7	Sequence 7, Appli
4	2235	99.2	428	12	US-10-286-926-3	Sequence 3, Appli
5	2235	99.2	428	12	US-10-286-926-5	Sequence 5, Appli
6	2235	99.2	428	12	US-10-231-913-127	Sequence 127, App
7	2235	99.2	428	14	US-10-091-085-3	Sequence 3, Appli
8	2235	99.2	428	14	US-10-091-085-5	Sequence 5, Appli
9	2235	99.2	428	14	US-10-092-063-3	Sequence 3, Appli
10	2235	99.2	428	14	US-10-092-063-5	Sequence 5, Appli
11	2089	92.8	405	12	US-10-286-926-25	Sequence 25, Appl
12	2089	92.8	405	14	US-10-092-063-25	Sequence 25, Appl
13	1979.5	87.9	427	12	US-10-231-913-126	Sequence 126, App
14	1822.5	80.9	465	14	US-10-092-063-39	Sequence 39, Appl
15	1645	73.0	330	9	US-09-925-299-876	Sequence 876, App
16	1645	73.0	330	11	US-09-925-299-876	Sequence 876, App
17	996	44.2	484	12	US-10-286-926-27	Sequence 27, Appl
18	996	44.2	484	12	US-10-231-913-123	Sequence 123, App
19	996	44.2	484	14	US-10-092-063-27	Sequence 27, Appl
20	993	44.1	467	12	US-10-231-913-36	Sequence 36, Appl
21	990.5	44.0	379	12	US-10-231-913-271	Sequence 271, App
22	989	43.9	484	12	US-10-231-913-124	Sequence 124, App
23	986	43.8	446	12	US-10-231-913-38	Sequence 38, Appl
24	983	43.7	455	12	US-10-231-913-125	Sequence 125, App
25	616.5	27.4	479	12	US-10-369-493-6447	Sequence 6447, Ap
26	605	26.9	476	9	US-09-835-147-4	Sequence 4, Appli
27	548	24.3	476	9	US-09-835-147-3	Sequence 3, Appli
28	535.5	23.8	556	12	US-10-369-493-2169	Sequence 2169, Ap
29	507.5	22.5	467	9	US-09-129-112-19	Sequence 19, Appl
30	502.5	22.3	459	9	US-09-129-112-9	Sequence 9, Appli
31	498	22.1	410	12	US-10-231-913-272	Sequence 272, App
32	493	21.9	462	9	US-09-129-112-2	Sequence 2, Appli
33	488.5	21.7	518	12	US-10-369-493-1713	Sequence 1713, Ap
34	486.5	21.6	462	9	US-09-129-112-15	Sequence 15, Appl
35	432	19.2	455	12	US-10-259-165-286	Sequence 286, App
36	426.5	18.9	459	12	US-10-259-165-74	Sequence 74, Appl
37	422	18.7	605	12	US-10-369-493-3333	Sequence 3333, Ap
38	382.5	17.0	485	12	US-10-369-493-6055	Sequence 6055, Ap
39	380	16.9	495	9	US-09-823-356-4	Sequence 4, Appli
40	374.5	16.6	557	12	US-10-369-493-6789	Sequence 6789, Ap
41	371.5	16.5	458	12	US-10-199-672-496	Sequence 496, App
42	371.5	16.5	458	12	US-10-187-749-496	Sequence 496, App
43	371.5	16.5	458	12	US-10-194-457-496	Sequence 496, App
44	371.5	16.5	458	12	US-10-184-642-496	Sequence 496, App
45	371.5	16.5	458	12	US-10-196-747-496	Sequence 496, App

ALIGNMENTS

RESULT 1
US-10-286-926-7
; Sequence 7, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457CON
; CURRENT APPLICATION NUMBER: US/10/286,926
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-926-7

```
  Query Match           100.0%;  Score 2252;  DB 12;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-218;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 2
US-10-091-085-7
; Sequence 7, Application US/10091085
; Publication No. US20020146772A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE: 28110/35761
;  CURRENT APPLICATION NUMBER: US/10/091,085
;  CURRENT FILING DATE:  2002-03-05
;  PRIOR APPLICATION NUMBER: 09/350,836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273,447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  PRIOR APPLICATION NUMBER: 09/122,449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/244,444
;  PRIOR FILING DATE: 1999-02-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-091-085-7

  Query Match           100.0%;  Score 2252;  DB 14;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-218;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 3
US-10-092-063-7
; Sequence 7, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
;  FILE REFERENCE: 28110/35908
;  CURRENT APPLICATION NUMBER: US/10/092,063
;  CURRENT FILING DATE:  2002-03-05
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350,836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273,447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/244,444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/122,449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-063-7

  Query Match           100.0%;  Score 2252;  DB 14;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-218;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
```

```
Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

```
RESULT 4
US-10-286-926-3
; Sequence 3, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  APPLICANT:  Yeung, George
;  TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 28110/36457CON
;  CURRENT APPLICATION NUMBER: US/10/286,926
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: 09/557,800
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/481,238
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 1999-08-09
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/122449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/244444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-286-926-3

  Query Match           99.2%;  Score 2235;  DB 12;  Length 428;
  Best Local Similarity  99.3%;  Pred. No. 8.5e-217;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

```
RESULT 5
US-10-286-926-5
; Sequence 5, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  APPLICANT:  Yeung, George
;  TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 28110/36457CON
;  CURRENT APPLICATION NUMBER: US/10/286,926
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: 09/557,800
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/481,238
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 1999-08-09
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/122449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/244444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-286-926-5

  Query Match           99.2%;  Score 2235;  DB 12;  Length 428;
  Best Local Similarity  99.3%;  Pred. No. 8.5e-217;
```

```
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 6
US-10-231-913-127
; Sequence 127, Application US/10231913
; Publication No. US20040005576A1
; GENERAL INFORMATION:
;  APPLICANT: Guo, Xiaojia S.
;  APPLICANT:  Li, Li
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shimkets, Richard A.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Alsobrook II, John P.
;  APPLICANT:  Edinger, Schlomit
;  APPLICANT:  Peyman, John A.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Boldog, Ference L.
;  APPLICANT:  Colman, Steven D.
;  APPLICANT:  Eisen, Andrew J.
;  APPLICANT:  Liu, Xiaohong
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Spaderna, Steven K.
;  APPLICANT:  Zerhusen, Bryan D.
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-216
;  CURRENT APPLICATION NUMBER: US/10/231,913
;  CURRENT FILING DATE:  2002-08-30
;  PRIOR APPLICATION NUMBER: 60/251,660
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: 60/255,029
;  PRIOR FILING DATE: 2000-12-12
;  PRIOR APPLICATION NUMBER: 60/260,326
;  PRIOR FILING DATE: 2001-01-08
;  PRIOR APPLICATION NUMBER: 60/263,800
;  PRIOR FILING DATE: 2001-01-24
;  PRIOR APPLICATION NUMBER: 60/269,942
;  PRIOR FILING DATE: 2001-02-20
;  PRIOR APPLICATION NUMBER: 60/286,183
;  PRIOR FILING DATE: 2001-04-24
;  PRIOR APPLICATION NUMBER: 60/313,627
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/318,712
;  PRIOR FILING DATE: 2001-09-12
;  NUMBER OF SEQ ID NOS: 292
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 127
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-231-913-127

  Query Match           99.2%;  Score 2235;  DB 12;  Length 428;
  Best Local Similarity  99.3%;  Pred. No. 8.5e-217;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 7
US-10-091-085-3
; Sequence 3, Application US/10091085
; Publication No. US20020146772A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE: 28110/35761
;  CURRENT APPLICATION NUMBER: US/10/091,085
;  CURRENT FILING DATE:  2002-03-05
;  PRIOR APPLICATION NUMBER: 09/350,836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273,447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
```

; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-091-085-3

Query Match 99.2%; Score 2235; DB 14; Length 428;
Best Local Similarity 99.3%; Pred. No. 8.5e-217;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 8
US-10-091-085-5
; Sequence 5, Application US/10091085
; Publication No. US20020146772A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/10/091,085
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-091-085-5

Query Match 99.2%; Score 2235; DB 14; Length 428;
Best Local Similarity 99.3%; Pred. No. 8.5e-217;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGILAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 9
US-10-092-063-3
; Sequence 3, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/10/092,063
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT

; ORGANISM: Homo sapiens
US-10-092-063-3

  Query Match 99.2%; Score 2235; DB 14; Length 428;
  Best Local Similarity 99.3%; Pred. No. 8.5e-217;
  Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 10
US-10-092-063-5
; Sequence 5, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/10/092,063
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-063-5

  Query Match 99.2%; Score 2235; DB 14; Length 428;
  Best Local Similarity 99.3%; Pred. No. 8.5e-217;
  Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 11
US-10-286-926-25
; Sequence 25, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457CON
; CURRENT APPLICATION NUMBER: US/10/286,926
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 405
; TYPE: PRT

; ORGANISM: Homo sapiens
US-10-286-926-25

Query Match 92.8%; Score 2089; DB 12; Length 405;
Best Local Similarity 99.2%; Pred. No. 4.4e-202;
Matches 397; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
              |||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
```

RESULT 12
US-10-092-063-25
; Sequence 25, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/10/092,063
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-063-25

Query Match 92.8%; Score 2089; DB 14; Length 405;
Best Local Similarity 99.2%; Pred. No. 4.4e-202;
Matches 397; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
              |||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
```

RESULT 13
US-10-231-913-126
; Sequence 126, Application US/10231913
; Publication No. US20040005576A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia S.
; APPLICANT: Li, Li
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Edinger, Schlomit
; APPLICANT: Peyman, John A.
; APPLICANT: Stone, David J.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Boldog, Ference L.
; APPLICANT: Colman, Steven D.
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Liu, Xiaohong
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-216
; CURRENT APPLICATION NUMBER: US/10/231,913
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 60/251,660
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: 60/255,029
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/260,326
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/263,800
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/269,942
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/286,183
; PRIOR FILING DATE: 2001-04-24

```
;  PRIOR APPLICATION NUMBER: 60/313,627
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/318,712
;  PRIOR FILING DATE: 2001-09-12
;  NUMBER OF SEQ ID NOS: 292
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 126
;   LENGTH: 427
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-231-913-126

  Query Match             87.9%;  Score 1979.5;  DB 12;  Length 427;
  Best Local Similarity   87.6%;  Pred. No. 5.5e-191;
  Matches  374;  Conservative   24;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              |||||| | |||:::|| | | :| |||||||:|||||||||||||| | ||||||||||||
Db          1 MATSWGAV-FMLIIACVGSTVFYREQQTWFEGVFLSSMCPINVSAGTFYGIMFDAGSTGT 59

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              |||||||||| |||| ||||:||||||||||||||||||||||| ||||||||||||||
Db         60 RIHVYTFVQKTAGQLPFLEGEIFDSVKPGLSAFVDQPKQGAETVQELLEVAKDSIPRSHW 119

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              ::|||||||||||||||| ||:||| ||:|||: |||||| |||||| |  ||| |||||
Db        120 ERTPVVLKATAGLRLLPEQKAQALLLEVEEIFKNSPFLVPDGSVSIMDGSYEGILAWVTV 179

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              |||||||||  |||||||||||||||||||||||||||||||||||||||||||:|||||
Db        180 NFLTGQLHGRGQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTFKLYTH 239

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              |||||||||||||||||| :|||||||||||||||||||||||||||||||||||:||
Db        240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGF 299

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||:|||||||||:  :|||||||||| || :|||||||:||||||||||
Db        300 EPCYAEVLRVVQGKLHQPEEVRGSAFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 359

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              |||||| :|:||||||||||:|||||||||||||| |:|||||||||||||||||||||
Db        360 EVCDNLGSFSSGSPFLCMDLTYITALLKDGFGFADGTLLQLTKKVNNIETGWALGATFHL 419

Qy        421 LQSLGIS 427
              |||||||
Db        420 LQSLGIT 426


RESULT 14
US-10-092-063-39
; Sequence 39, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT: Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
;  FILE REFERENCE: 28110/35908
;  CURRENT APPLICATION NUMBER: US/10/092,063
;  CURRENT FILING DATE:  2002-03-05
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350,836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273,447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/244,444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/122,449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
;   LENGTH: 465
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-092-063-39

  Query Match             80.9%;  Score 1822.5;  DB 14;  Length 465;
  Best Local Similarity   83.3%;  Pred. No. 4.5e-175;
  Matches  350;  Conservative   25;  Mismatches   42;  Indels    3;  Gaps    3;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              |||||| | |||:::|| | | :| |||||||:|||||||||||||| | ||||||||||||
Db          1 MATSWGAV-FMLIIACVGSTVFYREQQTWFEGVFLSSMCPINVSAGTFYGIMFDAGSTGT 59

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              |||||||||| |||| ||||:||||||||||||||||||||||| ||||||||||||||
Db         60 RIHVYTFVQKTAGQLPFLEGEIFDSVKPGLSAFVDQPKQGAETVQELLEVAKDSIPRSHW 119

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              ::|||||||||||||||| ||:||| ||:|||: |||||| |||||| |  ||| |||||
Db        120 ERTPVVLKATAGLRLLPEQKAQALLLEVEEIFKNSPFLVPDGSVSIMDGSYEGILAWVTV 179

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              |||||||||  |||||||||||||||||||||||||||||||||||||||||||:|||||
Db        180 NFLTGQLHGRGQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTFKLYTH 239

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              |||||||||||||||||| :|||||||||||||||||||||||||||||||||||:||
Db        240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGF 299

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||:|||||||||:  :|||||||||| || :|||||||:||||||||||
Db        300 EPCYAEVLRVVQGKLHQPEEVRGSAFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 359

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              |||||| :|:||||||||||:|||||||||||||| |:|||||||||||||||||||||
Db        360 EVCDNLGSFSSGSPFLCMDLTYITALLKDGLGFAERHPLTAHKE-SEQHRDW-LGLGGHL 417


RESULT 15
US-09-925-299-876
; Sequence 876, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 876
;   LENGTH: 330
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (97)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (106)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (124)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
```

```
;   LOCATION: (138)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (174)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (178)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (194)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-876

  Query Match             73.0%;  Score 1645;  DB 9;  Length 330;
  Best Local Similarity   96.3%;  Pred. No. 2.3e-157;
  Matches  316;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy        101 AETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVP 160
              |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ARAVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVP 62

Qy        161 KGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQT 220
              ||||||| | |||| |||||||||||||||| |||||||| ||||||||||||||||
Db         63 KGSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETXGTLDLGGAXTQITFLPQFEKTLEQT 122

Qy        221 PRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAE 280
              | ||||||||||||| |||||||||||||||||||||||||||||||| ||| ||||
Db        123 PXGYLTSFEMFNSTYXLYTHSYLGFGLKAARLATLGALETEGTDGHTFRSAXLPRXLEAE 182

Qy        281 WIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTD 340
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 WIFGGVKYQYGXNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTD 242

Qy        341 MIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 MIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 302

Qy        401 LTKKVNNIETGWALGATFHLLQSLGISH 428
              ||||||||||||||||||||||||||||
Db        303 LTKKVNNIETGWALGATFHLLQSLGISH 330


Search completed: January 29, 2004, 10:14:50
Job time : 30.5 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 10:05:35 ; Search time 29.5 Seconds
(without alignments)
3015.210 Million cell updates/sec

Title: US-10-091-085-3
Perfect score: 2250
Sequence: 1 MATSWGTVFFMLVVSCVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2250	100.0	428	12	US-10-286-926-3	Sequence 3, Appli
2	2250	100.0	428	12	US-10-286-926-5	Sequence 5, Appli
3	2250	100.0	428	12	US-10-231-913-127	Sequence 127, App
4	2250	100.0	428	14	US-10-091-085-3	Sequence 3, Appli
5	2250	100.0	428	14	US-10-091-085-5	Sequence 5, Appli
6	2250	100.0	428	14	US-10-092-063-3	Sequence 3, Appli
7	2250	100.0	428	14	US-10-092-063-5	Sequence 5, Appli
8	2235	99.3	428	12	US-10-286-926-7	Sequence 7, Appli
9	2235	99.3	428	14	US-10-091-085-7	Sequence 7, Appli
10	2235	99.3	428	14	US-10-092-063-7	Sequence 7, Appli
11	2104	93.5	405	12	US-10-286-926-25	Sequence 25, Appl
12	2104	93.5	405	14	US-10-092-063-25	Sequence 25, Appl
13	1994.5	88.6	427	12	US-10-231-913-126	Sequence 126, App
14	1837.5	81.7	465	14	US-10-092-063-39	Sequence 39, Appl
15	1660	73.8	330	9	US-09-925-299-876	Sequence 876, App
16	1660	73.8	330	11	US-09-925-299-876	Sequence 876, App
17	999	44.4	484	12	US-10-286-926-27	Sequence 27, Appl
18	999	44.4	484	12	US-10-231-913-123	Sequence 123, App
19	999	44.4	484	14	US-10-092-063-27	Sequence 27, Appl
20	996	44.3	467	12	US-10-231-913-36	Sequence 36, Appl
21	993.5	44.2	379	12	US-10-231-913-271	Sequence 271, App
22	992	44.1	484	12	US-10-231-913-124	Sequence 124, App
23	989	44.0	446	12	US-10-231-913-38	Sequence 38, Appl
24	986	43.8	455	12	US-10-231-913-125	Sequence 125, App
25	616.5	27.4	479	12	US-10-369-493-6447	Sequence 6447, Ap
26	590	26.2	476	9	US-09-835-147-4	Sequence 4, Appli
27	538.5	23.9	556	12	US-10-369-493-2169	Sequence 2169, Ap
28	533	23.7	476	9	US-09-835-147-3	Sequence 3, Appli
29	507.5	22.6	459	9	US-09-129-112-9	Sequence 9, Appli
30	503.5	22.4	467	9	US-09-129-112-19	Sequence 19, Appl
31	503	22.4	410	12	US-10-231-913-272	Sequence 272, App
32	498	22.1	462	9	US-09-129-112-2	Sequence 2, Appli
33	491.5	21.8	462	9	US-09-129-112-15	Sequence 15, Appl
34	483.5	21.5	518	12	US-10-369-493-1713	Sequence 1713, Ap
35	435	19.3	455	12	US-10-259-165-286	Sequence 286, App
36	429.5	19.1	459	12	US-10-259-165-74	Sequence 74, Appl
37	427	19.0	605	12	US-10-369-493-3333	Sequence 3333, Ap
38	387.5	17.2	485	12	US-10-369-493-6055	Sequence 6055, Ap
39	372.5	16.6	557	12	US-10-369-493-6789	Sequence 6789, Ap
40	372	16.5	630	12	US-10-369-493-1722	Sequence 1722, Ap
41	367	16.3	495	9	US-09-823-356-4	Sequence 4, Appli
42	358.5	15.9	458	12	US-10-199-672-496	Sequence 496, App
43	358.5	15.9	458	12	US-10-187-749-496	Sequence 496, App
44	358.5	15.9	458	12	US-10-194-457-496	Sequence 496, App
45	358.5	15.9	458	12	US-10-184-642-496	Sequence 496, App

ALIGNMENTS

RESULT 1
US-10-286-926-3
; Sequence 3, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457CON
; CURRENT APPLICATION NUMBER: US/10/286,926
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-926-3

```
  Query Match           100.0%;  Score 2250;  DB 12;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-215;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 2
US-10-286-926-5
; Sequence 5, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  APPLICANT:  Yeung, George
;  TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 28110/36457CON
;  CURRENT APPLICATION NUMBER: US/10/286,926
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: 09/557,800
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/481,238
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 1999-08-09
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/122449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/244444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-286-926-5

  Query Match           100.0%;  Score 2250;  DB 12;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-215;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 3
US-10-231-913-127
; Sequence 127, Application US/10231913
; Publication No. US20040005576A1
; GENERAL INFORMATION:
;  APPLICANT: Guo, Xiaojia S.
;  APPLICANT:  Li, Li
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shimkets, Richard A.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Alsobrook II, John P.
;  APPLICANT:  Edinger, Schlomit
;  APPLICANT:  Peyman, John A.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Boldog, Ference L.
;  APPLICANT:  Colman, Steven D.
;  APPLICANT:  Eisen, Andrew J.
;  APPLICANT:  Liu, Xiaohong
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Spaderna, Steven K.
;  APPLICANT:  Zerhusen, Bryan D.
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-216
;  CURRENT APPLICATION NUMBER: US/10/231,913
;  CURRENT FILING DATE:  2002-08-30
```

```
;  PRIOR APPLICATION NUMBER: 60/251,660
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: 60/255,029
;  PRIOR FILING DATE: 2000-12-12
;  PRIOR APPLICATION NUMBER: 60/260,326
;  PRIOR FILING DATE: 2001-01-08
;  PRIOR APPLICATION NUMBER: 60/263,800
;  PRIOR FILING DATE: 2001-01-24
;  PRIOR APPLICATION NUMBER: 60/269,942
;  PRIOR FILING DATE: 2001-02-20
;  PRIOR APPLICATION NUMBER: 60/286,183
;  PRIOR FILING DATE: 2001-04-24
;  PRIOR APPLICATION NUMBER: 60/313,627
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/318,712
;  PRIOR FILING DATE: 2001-09-12
;  NUMBER OF SEQ ID NOS: 292
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 127
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-231-913-127

  Query Match           100.0%;  Score 2250;  DB 12;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-215;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 4
US-10-091-085-3
; Sequence 3, Application US/10091085
; Publication No. US20020146772A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE: 28110/35761
;  CURRENT APPLICATION NUMBER: US/10/091,085
;  CURRENT FILING DATE:  2002-03-05
;  PRIOR APPLICATION NUMBER: 09/350,836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273,447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  PRIOR APPLICATION NUMBER: 09/122,449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/244,444
;  PRIOR FILING DATE: 1999-02-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-091-085-3

  Query Match           100.0%;  Score 2250;  DB 14;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-215;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 5
US-10-091-085-5
; Sequence 5, Application US/10091085
; Publication No. US20020146772A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE: 28110/35761
;  CURRENT APPLICATION NUMBER: US/10/091,085
;  CURRENT FILING DATE:  2002-03-05
;  PRIOR APPLICATION NUMBER: 09/350,836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273,447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
```

```
;  PRIOR APPLICATION NUMBER: 09/122,449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/244,444
;  PRIOR FILING DATE: 1999-02-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-091-085-5

  Query Match           100.0%;  Score 2250;  DB 14;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-215;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 6
US-10-092-063-3
; Sequence 3, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
;  FILE REFERENCE: 28110/35908
;  CURRENT APPLICATION NUMBER: US/10/092,063
;  CURRENT FILING DATE:  2002-03-05
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350,836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273,447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/244,444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/122,449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-063-3

  Query Match           100.0%;  Score 2250;  DB 14;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-215;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 7
US-10-092-063-5
; Sequence 5, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
;  FILE REFERENCE: 28110/35908
;  CURRENT APPLICATION NUMBER: US/10/092,063
;  CURRENT FILING DATE:  2002-03-05
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350,836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273,447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/244,444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/122,449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/118;205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
```

; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-063-5

Query Match 100.0%; Score 2250; DB 14; Length 428;
Best Local Similarity 100.0%; Pred. No. 2.3e-215;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 8
US-10-286-926-7
; Sequence 7, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457CON
; CURRENT APPLICATION NUMBER: US/10/286,926
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-926-7

Query Match 99.3%; Score 2235; DB 12; Length 428;
Best Local Similarity 99.3%; Pred. No. 7.3e-214;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              |||||||||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428
```

RESULT 9
US-10-091-085-7
; Sequence 7, Application US/10091085
; Publication No. US20020146772A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/10/091,085
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT

```
;  ORGANISM: Homo sapiens
US-10-091-085-7

  Query Match           99.3%;  Score 2235;  DB 14;  Length 428;
  Best Local Similarity  99.3%;  Pred. No. 7.3e-214;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              |||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 10
US-10-092-063-7
; Sequence 7, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
;  FILE REFERENCE: 28110/35908
;  CURRENT APPLICATION NUMBER: US/10/092,063
;  CURRENT FILING DATE:  2002-03-05
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350,836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273,447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/244,444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/122,449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-063-7
```

```
  Query Match           99.3%;  Score 2235;  DB 14;  Length 428;
  Best Local Similarity  99.3%;  Pred. No. 7.3e-214;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              |||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 11
US-10-286-926-25
; Sequence 25, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  APPLICANT:  Yeung, George
;  TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 28110/36457CON
;  CURRENT APPLICATION NUMBER: US/10/286,926
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: 09/557,800
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/481,238
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 1999-08-09
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/122449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/244444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
;   LENGTH: 405
;   TYPE: PRT
```

```
;   ORGANISM: Homo sapiens
US-10-286-926-25

  Query Match             93.5%;  Score 2104;  DB 12;  Length 405;
  Best Local Similarity   100.0%;  Pred. No. 7.7e-201;
  Matches  400;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
              ||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400


RESULT 12
US-10-092-063-25
; Sequence 25, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
;  FILE REFERENCE: 28110/35908
;  CURRENT APPLICATION NUMBER: US/10/092,063
;  CURRENT FILING DATE:  2002-03-05
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350,836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273,447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/244,444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/122,449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
;   LENGTH: 405
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-063-25

  Query Match             93.5%;  Score 2104;  DB 14;  Length 405;
  Best Local Similarity   100.0%;  Pred. No. 7.7e-201;
  Matches  400;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
              ||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400


RESULT 13
US-10-231-913-126
; Sequence 126, Application US/10231913
; Publication No. US20040005576A1
; GENERAL INFORMATION:
;  APPLICANT: Guo, Xiaojia S.
;  APPLICANT:  Li, Li
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shimkets, Richard A.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Alsobrook II, John P.
;  APPLICANT:  Edinger, Schlomit
;  APPLICANT:  Peyman, John A.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Boldog, Ference L.
;  APPLICANT:  Colman, Steven D.
;  APPLICANT:  Eisen, Andrew J.
;  APPLICANT:  Liu, Xiaohong
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Spaderna, Steven K.
;  APPLICANT:  Zerhusen, Bryan D.
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-216
;  CURRENT APPLICATION NUMBER: US/10/231,913
;  CURRENT FILING DATE:  2002-08-30
;  PRIOR APPLICATION NUMBER: 60/251,660
;  PRIOR FILING DATE: 2000-12-06
;  PRIOR APPLICATION NUMBER: 60/255,029
;  PRIOR FILING DATE: 2000-12-12
;  PRIOR APPLICATION NUMBER: 60/260,326
;  PRIOR FILING DATE: 2001-01-08
;  PRIOR APPLICATION NUMBER: 60/263,800
;  PRIOR FILING DATE: 2001-01-24
;  PRIOR APPLICATION NUMBER: 60/269,942
;  PRIOR FILING DATE: 2001-02-20
;  PRIOR APPLICATION NUMBER: 60/286,183
;  PRIOR FILING DATE: 2001-04-24
```

```
;  PRIOR APPLICATION NUMBER: 60/313,627
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/318,712
;  PRIOR FILING DATE: 2001-09-12
;  NUMBER OF SEQ ID NOS: 292
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 126
;   LENGTH: 427
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-231-913-126

  Query Match            88.6%;  Score 1994.5;  DB 12;  Length 427;
  Best Local Similarity  88.3%;  Pred. No. 6.9e-190;
  Matches  377;  Conservative   24;  Mismatches   25;  Indels    1;  Gaps    1;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              |||||| | |||:::|| | | :| ||||||||:|||||||||||||| | ||||||||||||
Db          1 MATSWGAV-FMLIIACVGSTVFYREQQTWFEGVFLSSMCPINVSAGTFYGIMFDAGSTGT 59

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              |||||||||  |||| ||||:|||||||||||||||||||||| ||||||||||||||
Db         60 RIHVYTFVQKTAGQLPFLEGEIFDSVKPGLSAFVDQPKQGAETVQELLEVAKDSIPRSHW 119

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ::|||||||||||||||| ||:||| ||:|||: |||||| |||||||||| |||||||||
Db        120 ERTPVVLKATAGLRLLPEQKAQALLLEVEEIFKNSPFLVPDGSVSIMDGSYEGILAWVTV 179

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              |||||||||  |||||||||||||||||||||||||||||||||||||||||||:|||||
Db        180 NFLTGQLHGRGQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTFKLYTH 239

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              |||||||||||||||||| :|||||||||||||||||||||||||||||||||||:||
Db        240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGF 299

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              |||||||||||:|||||||||:  :|||||||||| || :|||||||:||||||||||
Db        300 EPCYAEVLRVVQGKLHQPEEVRGSAFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 359

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              |||||| :|:||||||||||:|||||||||||||| |:|||||||||||||||||||||
Db        360 EVCDNLGSFSSGSPFLCMDLTYITALLKDGFGFADGTLLQLTKKVNNIETGWALGATFHL 419

Qy        421 LQSLGIS 427
              |||||||
Db        420 LQSLGIT 426


RESULT 14
US-10-092-063-39
; Sequence 39, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
;  APPLICANT: Ford, John
;  APPLICANT:  Mulero, Julio
;  TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
;  FILE REFERENCE: 28110/35908
;  CURRENT APPLICATION NUMBER: US/10/092,063
;  CURRENT FILING DATE:  2002-03-05
;  PRIOR APPLICATION NUMBER: 09/370,265
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: PCT/US99/16180
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: 09/350,836
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/273,447
;  PRIOR FILING DATE: 1999-03-19
;  PRIOR APPLICATION NUMBER: 09/244,444
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: 09/122,449
;  PRIOR FILING DATE: 1998-07-24
;  PRIOR APPLICATION NUMBER: 09/118,205
;  PRIOR FILING DATE: 1998-07-16
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
;   LENGTH: 465
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-092-063-39

  Query Match            81.7%;  Score 1837.5;  DB 14;  Length 465;
  Best Local Similarity  84.0%;  Pred. No. 3.5e-174;
  Matches  353;  Conservative   25;  Mismatches   39;  Indels    3;  Gaps    3;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              |||||| | |||:::|| | | :| ||||||||:|||||||||||||| | ||||||||||||
Db          1 MATSWGAV-FMLIIACVGSTVFYREQQTWFEGVFLSSMCPINVSAGTFYGIMFDAGSTGT 59

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              |||||||||  |||| ||||:|||||||||||||||||||||| ||||||||||||||
Db         60 RIHVYTFVQKTAGQLPFLEGEIFDSVKPGLSAFVDQPKQGAETVQELLEVAKDSIPRSHW 119

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180
              ::|||||||||||||||| ||:||| ||:|||: |||||| |||||||||| |||||||||
Db        120 ERTPVVLKATAGLRLLPEQKAQALLLEVEEIFKNSPFLVPDGSVSIMDGSYEGILAWVTV 179

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              |||||||||  |||||||||||||||||||||||||||||||||||||||||||:|||||
Db        180 NFLTGQLHGRGQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTFKLYTH 239

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              |||||||||||||||||| :|||||||||||||||||||||||||||||||||||:||
Db        240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGF 299

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              |||||||||||:|||||||||:  :|||||||||| || :|||||||:||||||||||
Db        300 EPCYAEVLRVVQGKLHQPEEVRGSAFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 359

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              |||||| :|:||||||||||:|||||||||||||| |:|||||||||||||||||||||
Db        360 EVCDNLGSFSSGSPFLCMDLTYITALLKDGLGFAERHPLTAHKE-SEQHRDW-LGLGGHL 417


RESULT 15
US-09-925-299-876
; Sequence 876, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 876
;   LENGTH: 330
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (97)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (106)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (124)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
```

; LOCATION: (138)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (174)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (178)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (194)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-876

Query Match 73.8%; Score 1660; DB 9; Length 330;
Best Local Similarity 97.3%; Pred. No. 1e-156;
Matches 319; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy        101 AETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVP 160
              |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ARAVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVP 62

Qy        161 KGSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQT 220
              |||||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||
Db         63 KGSVSIMDGSDEGILAWVTVNFLTGQLHGHRQETXGTLDLGGAXTQITFLPQFEKTLEQT 122

Qy        221 PRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAE 280
              | |||||||||||||| ||||||||||||||||||||||||||||||||| ||| ||||
Db        123 PXGYLTSFEMFNSTYXLYTHSYLGFGLKAARLATLGALETEGTDGHTFRSAXLPRXLEAE 182

Qy        281 WIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTD 340
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 WIFGGVKYQYGXNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTD 242

Qy        341 MIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 MIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 302

Qy        401 LTKKVNNIETGWALGATFHLLQSLGISH 428
              ||||||||||||||||||||||||||||
Db        303 LTKKVNNIETGWALGATFHLLQSLGISH 330
```

Search completed: January 29, 2004, 10:14:49
Job time : 30.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:35:09 ; Search time 9.5 Seconds
(without alignments)
2118.677 Million cell updates/sec

Title: US-10-091-085-7
Perfect score: 2252
Sequence: 1 MATSWGTVFFMLVVSCVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2235	99.2	428	1	ENP5_HUMAN	O75356 homo sapien
2	1979.5	87.9	427	1	ENP5_MOUSE	Q9wuz9 mus musculu
3	1918.5	85.2	469	1	ENP5_MESAU	Q9qyc8 mesocricetu
4	996	44.2	484	1	ENP6_HUMAN	O75354 homo sapien
5	983	43.7	455	1	ENP6_RAT	Q9er31 rattus norv
6	488.5	21.7	518	1	GDA1_YEAST	P32621 saccharomyc
7	484	21.5	455	1	NTPA_PEA	P52914 pisum sativ
8	441.5	19.6	454	1	APY_SOLTU	P80595 solanum tub
9	394	17.5	493	1	ENP1_CHICK	O93295 gallus gall
10	387	17.2	552	1	YBU4_CAEEL	Q21815 caenorhabdi
11	385	17.1	510	1	ENP1_MOUSE	P55772 mus musculu
12	382.5	17.0	485	1	YY4E_CAEEL	Q18411 caenorhabdi
13	375	16.7	513	1	ENP1_BOVIN	O18956 bos taurus
14	371	16.5	630	1	YND1_YEAST	P40009 saccharomyc
15	367	16.3	510	1	ENP1_PIG	Q9myu4 sus scrofa
16	366.5	16.3	613	1	ENP4_MOUSE	Q9dbt4 mus musculu
17	366	16.3	510	1	ENP1_HUMAN	P49961 homo sapien
18	363.5	16.1	511	1	ENP1_RAT	P97687 rattus norv
19	363	16.1	616	1	ENP4_HUMAN	Q9y227 homo sapien
20	357.5	15.9	529	1	ENP3_HUMAN	O75355 homo sapien
21	351.5	15.6	495	1	ENP2_HUMAN	Q9y513 homo sapien
22	321.5	14.3	495	1	ENP2_MOUSE	O55026 mus musculu
23	314.5	14.0	495	1	ENP2_RAT	O35795 rattus norv
24	310.5	13.8	494	1	ENP2_CHICK	P79784 gallus gall
25	150	6.7	628	1	NTP1_TOXGO	Q27893 toxoplasma
26	148	6.6	628	1	NTP2_TOXGO	Q27895 toxoplasma
27	141	6.3	592	1	NTP4_TOXGO	P52913 toxoplasma
28	105.5	4.7	553	1	VG1_SPV4	P11333 spiroplasma
29	104.5	4.6	535	1	DEXB_STRPN	Q54796 streptococc
30	100.5	4.5	634	1	SELB_MOOTH	Q46455 moorella th
31	100.5	4.5	1151	1	Y245_TREPA	O83273 treponema p
32	99	4.4	711	1	DPP3_YEAST	Q08225 saccharomyc
33	98.5	4.4	4427	1	PKSL_BACSU	Q05470 bacillus su
34	96.5	4.3	287	1	RM07_SCHPO	O14337 schizosacch
35	96	4.3	477	1	URT1_DESRO	P98119 desmodus ro
36	92	4.1	290	1	HEM6_CAUCR	Q9aat8 caulobacter
37	92	4.1	424	1	YI15_ARCFU	P56969 archaeoglob
38	92	4.1	1220	1	PTC1_BRARE	Q98864 brachydanio
39	92	4.1	1549	1	YIR3_YEAST	P40438 saccharomyc
40	92	4.1	1549	1	YJW2_YEAST	P40890 saccharomyc
41	91.5	4.1	796	1	TLR6_HUMAN	Q9y2c9 homo sapien
42	91	4.0	417	1	GLYA_BORBU	O51547 borrelia bu
43	91	4.0	946	1	ITH2_MOUSE	Q61703 mus musculu
44	90.5	4.0	279	1	HEM6_RICPR	Q9zc86 rickettsia
45	90.5	4.0	2076	1	FAS1_YARLI	P34229 y fatty aci

ALIGNMENTS

RESULT 1
ENP5_HUMAN
ID ENP5_HUMAN STANDARD; PRT; 428 AA.
AC O75356;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5 precursor
DE (EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like
DE 4) (ER-UDPase).
GN ENTPD5 OR CD39L4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukemia;
RX MEDLINE=98341119; PubMed=9676430;
RA Chadwick B.P., Frischauf A.-M.;
RT "The CD39-like gene family: identification of three new human members
RT (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of
RT the gene family from Drosophila melanogaster.";
RL Genomics 50:357-367(1998).
CC -!- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN
CC GLYCOPROTEINS FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC
CC RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER
CC NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE PYROPHOSPHATE
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC nucleotide + phosphate.
CC -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT LIVER, KIDNEY, PROSTATE,
CC TESTIS AND COLON. MUCH WEAKER EXPRESSION IN OTHER TISSUES.
CC -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------
DR EMBL; AF039918; AAC39885.1; -.
DR Genew; HGNC:3367; ENTPD5.
DR MIM; 603162; -.
DR GO; GO:0004002; F:adenosinetriphosphatase activity; TAS.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE_NEG.
KW Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW Endoplasmic reticulum; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 428 ECTONUCLEOSIDE TRIPHOSPHATE

```
FT                              DIPHOSPHOHYDROLASE 5.
FT   CARBOHYD     42     42       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    232    232       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    368    368       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   428 AA;  47517 MW;  830437A155DE4DDD CRC64;

  Query Match             99.2%;  Score 2235;  DB 1;  Length 428;
  Best Local Similarity   99.3%;  Pred. No. 5.2e-172;
  Matches  425;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |||||||||||||||||||||||||||||||||||||||| | |||| |||||
Db        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360

Qy        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420

Qy        421 LQSLGISH 428
              ||||||||
Db        421 LQSLGISH 428


RESULT 2
ENP5_MOUSE
ID   ENP5_MOUSE     STANDARD;      PRT;   427 AA.
AC   Q9WUZ9; O70214;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 5 precursor
DE   (EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like
DE   4) (ER-UDPase).
GN   ENTPD5 OR CD39L4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=98119025; PubMed=9457681;
RA   Chadwick B.P., Williamson J., Sheer D., Frischauf A.-M.;
RT   "cDNA cloning and chromosomal mapping of a mouse gene with homology to
RT   NTPases.";
RL   Mamm. Genome 9:162-164(1998).
RN   [2]
RP   SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND CHARACTERIZATION.
RC   TISSUE=Liver;
RX   MEDLINE=99298181; PubMed=10369669;
RA   Trombetta E.S., Helenius A.;
RT   "Glycoprotein reglucosylation and nucleotide sugar utilization in the
RT   secretory pathway: identification of a nucleoside diphosphatase in the
RT   endoplasmic reticulum.";
RL   EMBO J. 18:3282-3292(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN
CC       GLYCOPROTEINS FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC
CC       RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER
CC       NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE
CC       PYROPHOSPHATE.
CC   -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC       nucleotide + phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
CC   -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC   -!- TISSUE SPECIFICITY: Ubiquitous.
CC   -!- PTM: GLYCOSYLATED WITH HIGH MANNOSE N-LINKED GLYCANS.
CC   -!- MISCELLANEOUS: OPTIMAL PH IS NEUTRAL.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF006482; AAC05181.1; -.
DR   EMBL; AJ238636; CAB45533.1; -.
DR   EMBL; AK002618; BAB22234.1; -.
DR   MGD; MGI:1321385; Entpd5.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE_NEG.
KW   Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW   Endoplasmic reticulum; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    427       ECTONUCLEOSIDE TRIPHOSPHATE
FT                                DIPHOSPHOHYDROLASE 5.
FT   CARBOHYD     41     41       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    231    231       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    390    390       F -> L (IN REF. 1).
FT   CONFLICT    394    427       DGTLLQLTKKVNNIETGWALGATFHLLQSLGITS -> ERH
FT                                PLTAHKESEQHRDWLGLGGHLSPAPVSGHHQLRPSSTSEAC
FT                                ISEPVFSQEGVDSETFSDLSGKAWPETR (IN REF. 1).
SQ   SEQUENCE   427 AA;  47101 MW;  2F9DA2C342C55577 CRC64;

  Query Match             87.9%;  Score 1979.5;  DB 1;  Length 427;
  Best Local Similarity   87.6%;  Pred. No. 1.7e-151;
  Matches  374;  Conservative   24;  Mismatches   28;  Indels    1;  Gaps    1;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
```

```
Db          1 MATSWGAV-FMLIIACVGSTVFYREQQTWFEGVFLSSMCPINVSAGTFYGIMFDAGSTGT 59

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              |||||||||  |||| ||||:||||||||||||||||||||| ||||||||||||||
Db         60 RIHVYTFVQKTAGQLPFLEGEIFDSVKPGLSAFVDQPKQGAETVQELLEVAKDSIPRSHW 119

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              ::|||||||||||||||| ||:||| ||:|||: |||||| |||||| |  ||| |||||
Db        120 ERTPVVLKATAGLRLLPEQKAQALLLEVEEIFKNSPFLVPDGSVSIMDGSYEGILAWVTV 179

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQPEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              |||||||||  |||||||||||||||||||||||||||||||||||||||:|||||
Db        180 NFLTGQLHGRGQETVGTLDLGGASTQITFLPQPEKTLEQTPRGYLTSFEMFNSTFKLYTH 239

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              |||||||||||||||||| :|||||||||||||||||||||||||||||||||||||:||
Db        240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGF 299

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||:|||||||||:  :|||||||||||  || :|||||||:||||||||||
Db        300 EPCYAEVLRVVQGKLHQPEEVRGSAFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 359

Qy        361 EVCDNLENFTGSGPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              |||||| :|:||||||||||:|||||||||||||| |:||||||||||||||||||||||
Db        360 EVCDNLGSFSSGSPFLCMDLTYITALLKDGFGFADGTLLQLTKKVNNIETGWALGATFHL 419

Qy        421 LQSLGIS 427
              |||||||
Db        420 LQSLGIT 426


RESULT 3
ENP5_MESAU
ID   ENP5_MESAU     STANDARD;      PRT;   469 AA.
AC   Q9QYC8; Q9QYC9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 5 precursor
DE   (EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like
DE   4) (ER-UDPase) (Proto-oncogene cph).
GN   ENTPD5 OR CD39L4 OR CPH.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99142925; PubMed=9989819;
RA   Velasco J.A., Avila M.A., Notario V.;
RT   "The product of the cph oncogene is a truncated, nucleotide-binding
RT   protein that enhances cellular survival to stress.";
RL   Oncogene 18:689-701(1999).
CC   -!- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN
CC       GLYCOPROTEINS FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC
CC       RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER
CC       NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE PYROPHOSPHATE
CC       (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC       nucleotide + phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN FETAL CELLS AND MOST ADULT
CC       TISSUES.
CC   -!- DISEASE: CPH ONCOGENE HAS TRANSFORMING CAPACITY AND TUMORIGENIC
CC       POTENTIAL.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF084568; AAF22931.1; ALT_TERM.
DR   EMBL; AF084569; AAF22932.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE_NEG.
KW   Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW   Endoplasmic reticulum; Signal; Proto-oncogene.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    469       ECTONUCLEOSIDE TRIPHOSPHATE
FT                                DIPHOSPHOHYDROLASE 5.
FT   CARBOHYD     42     42       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    232    232       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   469 AA;  52125 MW;  03D8A23E0C73474B CRC64;

  Query Match             85.2%;  Score 1918.5;  DB 1;  Length 469;
  Best Local Similarity   86.0%;  Pred. No. 1.6e-146;
  Matches  361;  Conservative   22;  Mismatches   34;  Indels    3;  Gaps    1;

Qy          1 MATSWGTVFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ||| || |||:|:::|  | | :| |||||||:||||||| |||||| ||||||||||
Db          1 MATPWGAVFFLLMIACAGSTVFYREQQTWFEGVFLSSMCPANVSASTFYGIMFDAGSTGT 60

Qy         61 RIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
              |||||||||  |||| ||||:|||||||||||| |||||||:||  |||||||||||||
Db         61 RIHVYTFVQKAAGQLPFLEGEIFDSVKPGLSAFADQPKQGAQTVHALLEVAKDSIPRSHW 120

Qy        121 KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTV 180
              |:|||||||||||||| ||:||| ||:|||: ||||||  ||||| |  ||| |||||
Db        121 KRTPVVLKATAGLRLLPEQKAQALLLEVEEIFKMSPFLVPDDSVSIMDGSYEGILAWVTV 180

Qy        181 NFLTGQLHGHRQETVGTLDLGGASTQITFLPQPEKTLEQTPRGYLTSFEMFNSTYKLYTH 240
              |||||||||| |||:|||||||||||||||| |||||||| |||||||||:|||||
Db        181 NFLTGQLHGHSQETMGTLDLGGASTQITFLPQPSKTLEQTPRDYLTSFEMFNSTFKLYTH 240

Qy        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGF 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db        241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEMGF 300

Qy        301 EPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAR 360
              ||||||||||:|||||||::  |||||||||||| || :|||||||:||||||||||
Db        301 EPCYAEVLRVVQGKLHQPEEIRGSSFYAFSYYYDRAADTHLIDYEKGGVLKVEDFERKAR 360

Qy        361 EVCDNLENFTGSGPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHL 420
              |||||||:|||||||||||||||||||||||| |:||||||||||||  |    ||
Db        361 EVCDNLESFTSGSPFLCMDLSYITALLKDGFGFADGTLLQLTKKVNNIETGWPGG---HL 417


RESULT 4
ENP6_HUMAN
ID   ENP6_HUMAN     STANDARD;      PRT;   484 AA.
AC   O75354; Q9UJD1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 6 (EC 3.6.1.6)
DE   (NTPDase6) (CD39 antigen-like 2).
GN   ENTPD6 OR CD39L2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Keratinocytes;
RX   MEDLINE=98341119; PubMed=9676430;
RA   Chadwick B.P., Frischauf A.-M.;
RT   "The CD39-like gene family: identification of three new human members
RT   (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of
```

```
RT   the gene family from Drosophila melanogaster.";
RL   Genomics 50:357-367(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
CC   -!- FUNCTION: MIGHT SUPPORT GLYCOSYLATION REACTIONS IN THE GOLGI
CC       APPARATUS AND, WHEN RELEASED FROM CELLS, MIGHT CATALYZE THE
CC       HYDROLYSIS OF EXTRACELLULAR NUCLEOTIDES. HYDROLYZES PREFERENTIALLY
CC       NUCLEOSIDE 5'-DIPHOSPHATES, NUCLEOSIDE 5'-TRIPHOSPHATES ARE
CC       HYDROLYZED ONLY TO A MINOR EXTENT, THERE IS NO HYDROLYSIS OF
CC       NUCLEOSIDE 5'-MONOPHOSPHATES. THE ORDER OF ACTIVITY WITH DIFFERENT
CC       SUBSTRATES IS GDP > IDP >> UDP = CDP >> ADP (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC       nucleotide + phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI. BUT ALSO
CC       OCCURS IN A SOLUBLE EXTRACELLULAR FORM (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MOST TISSUES.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF039916; AAC39883.1; -.
DR   EMBL; AL035252; CAB41571.1; -.
DR   Genew; HGNC:3368; ENTPD6.
DR   MIM; 603160; -.
DR   GO; GO:0004002; F:adenosinetriphosphatase activity; TAS.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE_NEG.
KW   Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW   Signal-anchor; Golgi stack.
FT   DOMAIN        1     39       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     40     60       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       61    484       LUMENAL (POTENTIAL).
FT   CARBOHYD    220    220       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    284    284       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    138    138       V -> L (IN REF. 2).
FT   CONFLICT.   202    202       E -> K (IN REF. 2).
SQ   SEQUENCE   484 AA;  53233 MW;  27334B290DB8D64C CRC64;

  Query Match             44.2%;  Score 996;  DB 1;  Length 484;
  Best Local Similarity   52.4%;  Pred. No. 2e-72;
  Matches  204;  Conservative   54;  Mismatches  125;  Indels    6;  Gaps    4;

Qy         40 PINVSA---STLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQ 96
              |:  :|      |||||||||||||:||: |  : | : | |  | | :|||||||: |
Db         90 PLGTAADGHEVFYGIMFDAGSTGTRVHVFQFT-RPPRETPTLTHETFKAVKPGLSAYADD 148

Qy         97 PKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSP 156
                :: |: :: ||:|||  ||    || ||:|||||||||||||  ||: || :|||:|: ||
Db        149 VEKSAQGIRELLDVAKQDIPFDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASP 208

Qy        157 FLVPKGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKT 216
              |||    |||| | |||: ||:|:|||||  |      :|| ||||| |||| |||: | |
Db        209 FLVGDDCVSIMNGTDEGVSAWITINFLTGSLKTPGGSSVGMLDLGGGSTQIAFLPRVEGT 268

Qy        217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPR 275
              |: :| ||||:  ||| ||||||::|||| || :|||| || :| :   ||    | ||
Db        269 LQASPPGYLTALRMFNRTYKLYSYSYLGLGLMSARLAILGGVEGQPAKDGKELVSPCLSP 328

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDR 335
               : ||    | |:  | :      | | | |:: ::|: |||:   |||||||||
Db        329 SFKGEWEHAEVTYRVSGQKAAASLHELCAARVSEVLQNRVHRTEEVKHVDFYAFSYYYDL 388

Qy        336 AVDTDMIDYEKGGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395
              |    :|| |||| | | ||| |: ||  ||     ||| ||||:|:: ||:: |||
Db        389 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTLETQPQSSPFSCMDLTYVSLLLQE-FGFPR 447

Qy        396 STVLQLTKKVNNIETGWALGATFHLLQSL 424
              | ||:||:|::|:|| ||||| || : ||
Db        448 SKVLKLTRKIDNVETSWALGAIFHYIDSL 476


RESULT 5
ENP6_RAT
ID   ENP6_RAT       STANDARD;      PRT;   455 AA.
AC   Q9ER31;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 6 (EC 3.6.1.6)
DE   (NTPDase6) (CD39 antigen-like 2).
GN   ENTPD6 OR CD39L2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND CHARACTERIZATION.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=20498744; PubMed=11042118;
RA   Braun N., Fengler S., Ebeling C., Servos J., Zimmermann H.;
RT   "Sequencing, functional expression and characterization of rat
RT   NTPDase6, a nucleoside diphosphatase and novel member of the ecto-
RT   nucleoside triphosphate diphosphohydrolase family.";
RL   Biochem. J. 351:639-647(2000).
CC   -!- FUNCTION: MIGHT SUPPORT GLYCOSYLATION REACTIONS IN THE GOLGI
CC       APPARATUS AND, WHEN RELEASED FROM CELLS, MIGHT CATALYZE THE
CC       HYDROLYSIS OF EXTRACELLULAR NUCLEOTIDES. HYDROLYZES PREFERENTIALLY
CC       NUCLEOSIDE 5'-DIPHOSPHATES, NUCLEOSIDE 5'-TRIPHOSPHATES ARE
CC       HYDROLYZED ONLY TO A MINOR EXTENT, THERE IS NO HYDROLYSIS OF
CC       NUCLEOSIDE 5'-MONOPHOSPHATES. THE ORDER OF ACTIVITY WITH DIFFERENT
CC       SUBSTRATES IS GDP > IDP >> UDP = CDP >> ADP.
CC   -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC       nucleotide + phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI. BUT ALSO
CC       OCCURS IN A SOLUBLE EXTRACELLULAR FORM.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ277748; CAC16598.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE_NEG.
KW   Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW   Signal-anchor; Golgi stack.
FT   DOMAIN          1     12       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM       13     32       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                  (POTENTIAL).
FT   DOMAIN         33    455       LUMENAL (POTENTIAL).
FT   CARBOHYD      192    192       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      256    256       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   455 AA;  49899 MW;  19A22E8BAEF0F77B CRC64;

  Query Match             43.7%;  Score 983;  DB 1;  Length 455;
  Best Local Similarity   54.0%;  Pred. No. 2.1e-71;
  Matches  204;  Conservative   49;  Mismatches  119;  Indels    6;  Gaps    5;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLL 108
              ||||||||||||||||: |  : ||: | |   | | ::||||||: |  :: |: :| ||
Db         74 YGIMFDAGSTGTRIHVFQFA-RPPGETPTLTHETFKALKPGLSAYADDVEKSAQGIQELL 132

Qy        109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMT 168
               |||  ||    || ||:||||||||||||  ||: || :|||:|: |||||    ||||
Db        133 NVAKQHIPYDFWKATPLVLKATAGLRLLPGEKAQKLLQKVKEVFKASPFLVGDDCVSIMN 192

Qy        169 GQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSF 228
              | |||: ||:|||||||  |      :|| ||||| ||||||||: | ||: :| |:||:
Db        193 GTDEGVSAWITVNFLTGSLKTPGSSSVGMLDLGGGSTQITFLPRVEGTLQASPPGHLTAL 252

Qy        229 EMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPRWLEAEWIFGGVK 287
              :||| |:|||::|||| || :|||| || :| :   |      | ||      :|    |
Db        253 QMFNRTFKLYSYSYLGLGLMSARLAILGGVEGKPAEDDKELVSPCLSPRFRGKWEHAEVT 312

Qy        288 YQYGGNQEGEVG-FEPCYAEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEK 346
              |:  | :   || :| | : |  |:| |:|: || |   |||||||| |    :|| ||
Db        313 YRISGQK--AVGLYELCASRVSEVLRNKVHRTEEAQHVDFYAFSYYYDLAASFGLIDAEK 370

Qy        347 GGILKVEDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVN 406
              || | |||  |: ||  ||     ||| ||||:||: || : |||    ||:| :|::
Db        371 GGSLVVGDFEIAAKYVCRTLETQPPSSPFACMDLTYISLLLHE-FGFPGDKVLKLARKID 429

Qy        407 NIETGWALGATFHLLQSL 424
              |:|| ||||| || : ||
Db        430 NVETSWALGAIFHYIDSL 447


RESULT 6
GDA1_YEAST
ID   GDA1_YEAST     STANDARD;      PRT;   518 AA.
AC   P32621;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Guanosine-diphosphatase (EC 3.6.1.42) (GDPase).
GN   GDA1 OR YEL042W OR SYGP-ORF16.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=G2-9;
RX   MEDLINE=93308137; PubMed=8391537;
RA   Abeijon C., Yanagisawa K., Mandon E.C., Haeusler A., Moremen K.,
RA   Hirschberg C.B., Robbins P.W.;
RT   "Guanosine diphosphatase is required for protein and sphingolipid
RT   glycosylation in the Golgi lumen of Saccharomyces cerevisiae.";
RL   J. Cell Biol. 122:307-323(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   PubMed=9169868;
RA   Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA   Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA   Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA   Hunicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA   Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA   Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA   Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
RL   Nature 387:78-81(1997).
CC   -!- FUNCTION: AFTER TRANSFER OF SUGARS TO ENDOGENOUS MACROMOLECULAR
CC       ACCEPTORS, THE ENZYME CONVERTS NUCLEOSIDE DIPHOSPHATES TO
CC       NUCLEOSIDE MONOPHOSPHATES WHICH IN TURN EXIT THE GOLGI LUMEN IN
CC       A COUPLED ANTIPORTER REACTION, ALLOWING ENTRY OF ADDITIONAL
CC       NUCLEOTIDE SUGAR FROM THE CYTOSOL.
CC   -!- CATALYTIC ACTIVITY: GDP + H(2)O = GMP + phosphate.
CC   -!- PATHWAY: Glycosylation.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L19560; AAA34656.1; -.
DR   EMBL; U18779; AAB65000.1; -.
DR   PIR; A40732; A40732.
DR   SGD; S0000768; GDA1.
DR   GO; GO:0004382; F:guanosine diphosphatase activity; IDA.
DR   GO; GO:0045134; F:uridine diphosphatase activity; IDA.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Golgi stack; Glycoprotein; Transmembrane; Signal-anchor.
FT   DOMAIN          1      9       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM       10     24       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                  (POTENTIAL).
FT   DOMAIN         25    518       LUMENAL (POTENTIAL).
FT   CARBOHYD       41     41       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      280    280       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      335    335       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   518 AA;  56821 MW;  9A61182D5ED22ADF CRC64;

  Query Match             21.7%;  Score 488.5;  DB 1;  Length 518;
  Best Local Similarity   33.0%;  Pred. No. 1.3e-31;
  Matches  143;  Conservative   56;  Mismatches  159;  Indels   75;  Gaps   15;

Qy         49 YGIMFDAGSTGTRIHVYTF-VQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGL 107
              | || ||||||:|:|:| | |   |   | |  | || ::||||:|       | ::  |
Db         93 YVIMIDAGSTGSRVHIYKFDVCTSP---PTLLDEKFDMLEPGLSSFDTDSVGAANSLDPL 149

Qy        108 LEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKS-PFLVPKG-SVS 165
              |:|| : :|      ||| :||||||||| : |:  :|  |::   |  || | :|  ||
Db        150 LKVAMNYVPIKARSCTPVAVKATAGLRLLGDAKSSKILSAVRDHLEKDYPFPVVEGDGVS 209

Qy        166 IMTGQDEGIFAWVTVNFLTGQL--HGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRG 223
              || | :||:|||:| |:| | :  :| :  |    |||| |||| | | |    :
Db        210 IMGGDEEGVFAWITTNYLLGNIGANGPKLPTAAVFDLGGGSTQIVFEPTFPINEKMVDGE 269

Qy        224 YLTSFEMFNSTYKLYTHSYLGFGLKAAR----------LATLGALETEGTDGHTFRSAC 272
              :   :  :  | ||  |:||:||| |            |     |: : |  |  | |
Db        270 HKFDLKFGDENYTLYQFSHLGYGLKEGRNKVNSVLVENALKDGKILKGDNTKTHQLSSPC 329
```

```
Qy        273 LPRWLEA--EWI-----------FGGVKYQYGG----------NQEGEVGFEPCYAEVLR 309
              ||  : |  | :          | |      |          |:: :    ||
Db        330 LPPKVNATNEKVTLESKETYTIDFIGPDEPSGAQCRFLTDEILNKDAQCQSPPCSF---- 385

Qy        310 VVRGKLHQPEEV----QRGSFYAFSYYYDRAVDTDM-IDYEKGGILKVEDFERKAREVC- 363
                :|||  |    :    | |||:|||     | : :       : :    || ||
Db        386 ---NGVHQPSLVRTFKESNDIYIFSYFYDRTRPLGMPLSF------TLNELNDLARIVCK 436

Qy        364 ------------DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIET 410
                         :|:   | | | |:|||:  :||  |:       |:  ||: | |
Db        437 GEETWNSVFSGIAGSLDELESDSHF-CLDLSFQVSLLHTGYDIPLQRELRTGKKIANKEI 495

Qy        411 GWALGATFHLLQS 423
              || |||:  ||:: .
Db        496 GWCLGASLPLLKA 508
```

RESULT 7
NTPA_PEA
ID NTPA_PEA STANDARD; PRT; 455 AA.
AC P52914;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nucleoside-triphosphatase (EC 3.6.1.15) (Nucleoside triphosphate
DE phosphohydrolase) (NTPase) (Apyrase).
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Alaska; TISSUE=Plumule;
RX MEDLINE=96197404; PubMed=8616230;
RA Hsieh H., Tong C.G., Thomas C., Roux S.J.;
RT "Light-modulated abundance of an mRNA encoding a calmodulin-regulated,
RT chromatin-associated NTPase in pea.";
RL Plant Mol. Biol. 30:135-147(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Alaska; TISSUE=Stem;
RA Shibata K., Abe S., Davies E.;
RT "Structure of the coding region and mRNA variants of the apyrase from
RT Pisum sativum.";
RL Acta Physiol. Plant. 20:3-13(2001).
CC -!- FUNCTION: MIGHT BE INVOLVED IN RNA TRANSPORT OUT OF NUCLEI.
CC -!- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + phosphate.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; Z32743; CAA83655.1; -.
DR EMBL; AB027614; BAB18900.1; -.
DR PIR; S65147; S48859.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW Hydrolase; Nuclear protein.
SQ SEQUENCE 455 AA; 50072 MW; 50FDF0023ABC4299 CRC64;

Query Match 21.5%; Score 484; DB 1; Length 455;
Best Local Similarity 33.4%; Pred. No. 2.5e-31;
Matches 139; Conservative 62; Mismatches 161; Indels 54; Gaps 14;

```
Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEG-EVFDSVKPGLSAFVDQPKQGAETVQGL 107
              | ::|||||||:|||| | | :   | | :| | :: : ||||:: : |:| |::  |
Db         44 YAVVFDAGSTGSRIHVYHFNQNL-DLLHIGKGVEYYNKITPGLSSYANNPEQAAKSLIPL 102

Qy        108 LEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIF-RKSPFLVPKGSVSI 166
              || |:| :|     |||| | ||||||||    :: :|  |:::   :| | |   :|||
Db        103 LEQAEDVVPDDLQPKTPVRLGATAGLRLLNGDASEKILQSVRDMLSNRSTFNVQPDAVSI 162

Qy        167 MTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPR---- 222
              : |  || : |||||:  | |      :||| :|||| | |: :   :|| :  |:
Db        163 IDGTQEGSYLWVTVNYALGNLGKKYTKTVGVIDLGGGSVQMAYAVS-KKTAKNAPKVADG 221

Qy        223 --GYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRS--ACLPRWLE 278
                |:    :    | || |||| || :|:|  |             | ||   ||
Db        222 DDPYIKKVVLKGIPYDLYVHSYLHFGREASRAEILKL---------TPRSPNPCLLAGFN 272

Qy        279 AEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSF---------- 326
              : : | :::      |   | |   : : :  ||: |   |  :|
Db        273 GIYTYSGEEFKATAYTSG-ANFNKCKNTIRKAL--KLNYPCPYQNCTFGGIWNGGGGNGQ 329

Qy        327 ---YAFSYYYDRAVDTDMIDYEKGG-ILKVEDFERKAREVCD-NLENFTSGSPFL----- 376
                :| | ::    || |:|      ||:  | | ||:| |  | |:  |  |||
Db        330 KNLFASSSFFYLPEDTGMVDASTPNFILRPVDIETKAKEACALNFEDAKSTYPFLDKKNV 389

Qy        377 ----CMDLSYITALLKDGFGFADSTVLQLTKKVNN----IETGWALGATFHLLQSL 424
                 |||| |   || ||||     :   |::     :|  | ||     : :|
Db        390 ASYVCMDLIYQYVLLVDGFGLDPLQKITSGKEIEYQDAIVEAAWPLGNAVEAISAL 445
```

RESULT 8
APY_SOLTU
ID APY_SOLTU STANDARD; PRT; 454 AA.
AC P80595; Q43164;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Apyrase precursor (EC 3.6.1.5) (ATP-diphosphatase) (Adenosine
DE diphosphatase) (ADPase) (ATP-diphosphohydrolase).
GN RROP1.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 59-160; 236-253 AND 332-345.
RC TISSUE=Tuber;
RX MEDLINE=96158985; PubMed=8579614;
RA Handa M., Guidotti G.;
RT "Purification and cloning of a soluble ATP-diphosphohydrolase
RT (apyrase) from potato tubers (Solanum tuberosum).";
RL Biochem. Biophys. Res. Commun. 218:916-923(1996).
RN [2]
RP SEQUENCE OF 42-54; 68-95 AND 236-253.
RC STRAIN=cv. Desiree;
RX MEDLINE=96355615; PubMed=8703025;
RA Vasconcelos E.G., Ferreira S.T., de Carvalho T.M.U., de Souza W.,
RA Kettlun A.M., Mancilla M., Valenzuela M.A., Verjovski-Almeida S.;
RT "Partial purification and immunohistochemical localization of ATP
RT diphosphohydrolase from Schistosoma mansoni. Immunological cross-
RT reactivities with potato apyrase and Toxoplasma gondii nucleoside
RT triphosphate hydrolase.";
RL J. Biol. Chem. 271:22139-22145(1996).
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF PHOSPHOANHYDRIDE BONDS OF
CC NUCLEOSIDE TRI- AND DI-PHOSPHATES.
CC -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC -!- COFACTOR: CALCIUM.
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Probable).
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; U58597; AAB02720.1; -.
DR   PIR; JC4616; JC4616.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Transmembrane; Calcium; Signal.
FT   SIGNAL       1     30       POTENTIAL.
FT   CHAIN       31    454       APYRASE.
FT   TRANSMEM    426   446       POTENTIAL.
FT   CARBOHYD    151   151       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    262   262       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   454 AA;  50041 MW;  9D9EFE431DA2F52F CRC64;

  Query Match             19.6%;  Score 441.5;  DB 1;  Length 454;
  Best Local Similarity   28.6%;  Pred. No. 6.4e-28;
  Matches  130;  Conservative   78;  Mismatches  178;  Indels   69;  Gaps   16;

Qy         25 NQQTWFEGIFLS-------SMCPINVSA-----------STLYGIMFDAGSTGTRIHVY 65
              || : |  | |:       |:   ||:|             |  | ::|||||||:|:||:
Db          3 NQNSHFIFIILAIFLVLPLSLLSKNVNAQIPLRRHLLSHESEHYAVIFDAGSTGSRVHVF 62

Qy         66 TFVQKMPGQLPILEG-EVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTP 124
               | :|: | |||    | | : :||||:: : ||  | ::: ||: |:  :|:    :||
Db         63 RFDEKL-GLLPIGNNIEYFMATEPGLSSYAEDPKAAANSLEPLLDGAEGVVPQELQSETP 121

Qy        125 VVLKATAGLRLLPEHKAKALLFEVKEIFR-KSPFLVPKGSVSIMTGQDEGIFAWVTVNFL 183
               : | ||||||:|    |: :|  |: : : :| |       |:|: |  || : |  :|:|
Db        122 LELGATAGLRMLKGDAAEKILQAVRNLVKNQSTFHSKDQWVTILDGTQEGSYMWAAINYL 181

Qy        184 TGQLHGHRQETVGTLDLGGASTQITFL---PQFEKTLE-QTPRGYLTSFEMFNSTYKLYT 239
               | |    : |  |:|||| | |: :     || |  : :    |:    : :  | ||
Db        182 LGNLGKDYKSTTATIDLGGGSVQMAYAISNEQFAKAPQNEDGEPYVQQKHLMSKDYNLYV 241

Qy        240 HSYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVG 299
              |||| :|  | |     |  |        : |    :  : :||| |:    ::|
Db        242 HSYLNYGQLAGRAEIFKASRNES-------NPCALEGCDGYYSYGGVDYKVKAPKKGS-- 292

Qy        300 FEPCYAEVLRVVRG--KLHQPEEVQRGSF--------------YAFSYYYDRAVDTDMI 342
                  :   |: |   |::    ::  :|               :| |::||      ::
Db        293 ---SWKRCRRLTRHALKINAKCNIEECTFNGVWNGGGGDGQKNIHASSFFYDIGAQVGIV 349

Qy        343 DYE-KGGILKVEDFERKAREVCD-NLENFTS--------GSPFLCMDLSYITALLKDGFG 392
              | :    : |   :   |:  |  |: :  |          |:|||| |   || ||||
Db        350 DTKFPSALAKPIQYLNAAKVACQTNVADIKSIFPKTQDRNIPYLCMDLIYEYTLLVDGFG 409

Qy        393 F---ADSTVLQLTKKVNN-IETGWALGATFHLLQS 423
                : ||:   :  |  :   | ||   .|: |
Db        410 LNPHKEITVIHDVQYKNYLVGAAWPLGCAIDLVSS 444


RESULT 9
ENP1_CHICK
ID   ENP1_CHICK     STANDARD;      PRT;   493 AA.
AC   O93295;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE   (NTPDase1) (Ecto-ATP diphosphohydrolase) (ATPDase) (Lymphoid cell
DE   activation antigen) (Ecto-apyrase) (CD39 antigen).
GN   ENTPD1 OR CD39.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-21 AND 150-156.
RC   TISSUE=Oviduct;
RX   MEDLINE=98298108; PubMed=9632655;
RA   Nagy A.K., Knowles A.F., Nagami G.T.;
RT   "Molecular cloning of the chicken oviduct ecto-ATP-
RT   diphosphohydrolase.";
RL   J. Biol. Chem. 273:16043-16049(1998).
RN   [2]
RP   SEQUENCE OF 1-17.
RC   TISSUE=Stomach;
RX   MEDLINE=97442428; PubMed=9295305;
RA   Lewis-Carl S., Kirley T.L.;
RT   "Immunolocalization of the ecto-ATPase and ecto-apyrase in chicken
RT   gizzard and stomach. Purification and N-terminal sequence of the
RT   stomach ecto-apyrase.";
RL   J. Biol. Chem. 272:23645-23652(1997).
CC   -!- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC       nucleotides to regulate purinergic neurotransmission. Could also
CC       be implicated in the prevention of platelet aggregation.
CC       Hydrolyzes ATP and ADP equally well (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- PTM: N-GLYCOSYLATED.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; AF041355; AAC26491.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT   DOMAIN       1      7       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     8     28       POTENTIAL.
FT   DOMAIN      29    463       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    464   486       POTENTIAL.
FT   DOMAIN      487   493       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     65    65       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     79    79       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    133   133       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    223   223       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234   234       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    267   267       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    324   324       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    330   330       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    361   361       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    372   372       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    382   382       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    445   445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     16    16       C -> W (IN REF. 2).
FT   CONFLICT     21    21       I -> G (IN REF. 1; AA SEQUENCE).
SQ   SEQUENCE   493 AA;  54034 MW;  F14FF4C3AA2F3603 CRC64;

  Query Match             17.5%;  Score 394;  DB 1;  Length 493;
  Best Local Similarity   27.8%;  Pred. No. 4.7e-24;
  Matches  133;  Conservative   65;  Mismatches  169;  Indels  112;  Gaps   18;

Qy         11 MLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYTF-VQ 69
              :|  :|| | ::         : :||           | ||::|||||| | ::|| :
Db         11 LLTATCVFSIIALILSAVDVKDVFL--------PPGTKYGLVFDAGSTHTALYVYQWPAD 62

Qy         70 KMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKA 129
              |  |   : : |       |:|:: | |     :::   |: |   ||     :||  | |
Db         63 KENGTGIVSQVESCTVNGSGISSYADDPAGAGASLKPCLDKAMAVIPVEQQWQTPTYLGA 122

Qy        130 TAGLRLLPEH---KAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTVNFLTGQ 186
```

```
Db        123 TAGMRLLREQNSTKAEQVFAEVSKAIREFP--VDFRGAQILTGNEEGSFGWITVNYLLET 180

Qy        187 L----------HGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYK 236

Db        181 LIKFSFAGKWEHPQNTEVLGALDLGGASTQITFQPGV--TIEDKNTSVL--FRLYGTNYS 236

Qy        237 LYTHSYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQE- 295

Db        237 LYTHSYLCYGQIQASKRLMAALHQDGSYVQNISHPCYPK----------------GYRRI 280

Qy        296 ---GEVGFEPCY--------AEVLRVVR--------------------------GKLHQ 317

Db        281 ITIAEIYDSPCVPTPSMLSPAQILTVTGTGNPAACPTAILKLFNLTCGANRTCGFDGVYQ 340

Qy        318 PEEVQRGSFYAFS-YYYDRA------------VDTDMIDY-EKGGILKVEDFERKAREVC 363

Db        341 PP--VRGQFFAFAGFYYTFSFLNLTGQQSLSHVNATVWDFCNKNWSELVETFPQNK---- 394

Qy        364 DNLENFTSGSPFLCMDLSYITALLKDGFGFADST--VLQLTKKVNNIETGWALGATFHL 420

Db        395 EHLHTY-------CVVGLYILTLLVDGYKFDEHTWSNIHFSQKAGNADIGWTLGFMLNL 446


RESULT 10
YBU4_CAEEL
ID   YBU4_CAEEL     STANDARD;      PRT;   552 AA.
AC   Q21815;
DT   01-NOV-1997 (Rel. 35, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein R07E4.4 in chromosome X.
GN   R07E4.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Miller N.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   REVISIONS.
RA   Waterston R.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U39652; AAA80403.2; -.
DR   WormPep; R07E4.4; CE28748.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hypothetical protein; Transmembrane; Hydrolase.
FT   TRANSMEM      7     27       POTENTIAL.
FT   TRANSMEM    490    510       POTENTIAL.
SQ   SEQUENCE   552 AA;  62510 MW;  E5DC32C858AE4D94 CRC64;

  Query Match             17.2%;  Score 387;  DB 1;  Length 552;
  Best Local Similarity   28.6%;  Pred. No. 2e-23;
  Matches  126;  Conservative   70;  Mismatches  159;  Indels   86;  Gaps   20;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDS------VKPGLSAFVDQPKQGAE 102

Db         44 YGVICDAGSTGTRLFVYNWISTSDSELIQIEPVIYDNKPVMKKISPGLSTFGTKPAQAAE 103

Qy        103 TVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFL---- 158

Db        104 YLRPLMELAERHIPEEKRPYTPVFIFATAGMRLIPDEQKEAVL---KNLRNKLPKITSMQ 160

Qy        159 VPKGSVSIMTGQDEGIFAWVTVNFLTGQLH-----------GH-RQETVGTLDLGGAST 205

Db        161 VLKEHIRIIEGKWEGIYSWIAVNYALGKFNKTATLDFPGTSPAHARQKTVGMIDMGGASA 220

Qy        206 QITF-LPQFEK----TLEQTPRGYLTSFEMFNSTYKLYTHSYLGF----GLKAARLATLG 256

Db        221 QIAFELPDTDSFSSINVENINLGCREDDSLFK--YKLFVTTFLGYGVNEGIRKYEHMLLS 278

Qy        257 ALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGN--QEG----------------E 297

Db        279 KLKDQ--NGTVIQDDCMPLNLHKT-----VTLENGENFVRRGTGNWNTCSNEVKKLLNPE 331

Qy        298 VGFEPCYAEVLRVVRGKLHQPE-EVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFE 356

Db        332 SSSEVCKAEAAKCYFGAVPAPSIPLSNIEMYGFSEYWYSTHDV----LGLGGQYDAENIA 387

Qy        357 RKAREVCD------NLENFTSGSP--------FLCMDLSYITALLKDGFGFADST--VLQ 400

Db        388 KKTQQYCSKRWSTIQAESKKQLYPRADEERLRTQCFKSAWITSVLHDGFS-VDKTHNKFQ 446

Qy        401 LTKKVNNIETGWALGA-TFHL 420

Db        447 SVSTIAGQEVQWALGAMIYHM 467


RESULT 11
ENP1_MOUSE
ID   ENP1_MOUSE     STANDARD;      PRT;   510 AA.
AC   P55772;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE   (NTPDase1) (Ecto-ATP diphosphohydrolase) (ATPDase) (Lymphoid cell
DE   activation antigen) (Ecto-apyrase) (CD39 antigen).
GN   ENTPD1 OR CD39.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95015846; PubMed=7930580;
RA   Maliszewski C.R., Delespesse G.J.T., Schoenborn M.A., Armitage R.J.,
RA   Fanslow W.C., Nakajima T., Baker E., Sutherland G.R., Poindexter K.,
RA   Birks C., Alpert A., Friend D., Gimpel S.D., Gayle R.B. III;
RT   "The CD39 lymphoid cell activation antigen. Molecular cloning and
RT   structural characterization.";
RL   J. Immunol. 153:3574-3583(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=98399871; PubMed=9730622;
RA   Schoenborn M.A., Jenkins N.A., Copeland N.G., Gilbert D.J.,
RA   Gayle R.B. III, Maliszewski C.R.;
RT   "Gene structure and chromosome location of mouse Cd39 coding for an
RT   ecto-apyrase.";
RL   Cytogenet. Cell Genet. 81:287-289(1998).
CC   -!- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC       nucleotides to regulate purinergic neurotransmission. Could also
CC       be implicated in the prevention of platelet aggregation.
CC       Hydrolyzes ATP and ADP equally well.
CC   -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC   -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF037366; AAB92259.1; -.
DR   EMBL; AF041818; AAC83203.1; -.
DR   EMBL; AF041812; AAC83203.1; JOINED.
DR   EMBL; AF041813; AAC83203.1; JOINED.
DR   EMBL; AF041814; AAC83203.1; JOINED.
DR   EMBL; AF041815; AAC83203.1; JOINED.
DR   EMBL; AF041816; AAC83203.1; JOINED.
DR   EMBL; AF041817; AAC83203.1; JOINED.
DR   MGD; MGI:102805; Entpd1.
DR   GO; GO:0005605; C:basal lamina; IDA.
DR   GO; GO:0004050; F:apyrase activity; IDA.
DR   GO; GO:0006200; P:ATP catabolism; IDA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IDA.
DR   GO; GO:0030168; P:platelet activation; IDA.
DR   GO; GO:0009181; P:purine ribonucleoside diphosphate catabolism; IDA.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT   DOMAIN        1     16       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     17     37       POTENTIAL.
FT   DOMAIN       38    478       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    479    499       POTENTIAL.
FT   DOMAIN      500    510       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     73     73       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    291    291       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    428    428       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    457    457       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   510 AA;  57205 MW;  8E6A6113D2E13930 CRC64;

  Query Match             17.1%;  Score 385;  DB 1;  Length 510;
  Best Local Similarity   27.5%;  Pred. No. 2.6e-23;
  Matches  123;  Conservative   72;  Mismatches  170;  Indels   82;  Gaps   19;

Qy         32 GIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVK-PGL 90
              |:  :   | ||     |||: ||||: | :::| :  :        :::      || ||:
Db         36 GLTQNKPLPENVK----YGIVLDAGSSHTNLYIYKWPAEKENDTGVVQQLEECQVKGPGI 91

Qy         91 SAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLL---PEHKAKALLFE 147
              | :  :  :    :   :|:: : || |   :||| | ||||:|||    |  |  :|
Db         92 SKYAQKTDEIGAYLAECMELSTELIPTSKHHQTPVYLGATAGMRLLRMESEQSADEVLAA 151

Qy        148 VKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTVNFLTGQL----------HGHRQETVG 196
              |    :  ||        |:|||:|| : |:|:|:| |:               :||| |
Db        152 VSTSLKSYPF--DFQGAKIITGQEEGAYGWITINYLLGRFTQEQSWLSLISDSQKQETFG 209

Qy        197 TLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLG 256
              ||||||||||||:||   |:| :|   |  | ::   | :||||:| :|   |    |
Db        210 ALDLGGASTQITFVPQ-NSTIE-SPENSL-QFRLYGEDYTVYTHSFLCYGKDQALWQKL- 265

Qy        257 ALETEGTDGHTFRSACL-----------------PRWLEAEWIFGGVKYQYGGNQEGEV 298
              | : : |   :  |                   :  | :  |   : |  |:
Db        266 AKDIQVSSGGVLKDPCFNPGYEKVVNVSELYGTPCTKRFEKKLPFDQFRIQGTGD----- 320

Qy        299 GFEPCYAEVLRVVRGKLHQPEE----------VQRGSFYAFSYYYDRAVDTDMIDY---- 344
               :| |:  :| :      | |               ||| ||| :|       ::|:
Db        321 -YEQCHQSILELFNNS-HCPYSQCAFNGVFLPPLHGSFGAFSAFY------FVMDFFKKV 372

Qy        345 EKGGILKVEDFERKAREVCD-NLENFTSGSPFL--------CMDLSYITALLKDGFGFAD 395
              |  ::  |      :  |  : |   :  | :        |   :|| :||: |: | |
Db        373 AKNSVISQEKMTEITKNFCSKSWEETKTSYPSVKEKYLSEYCFSGAYILSLLQ-GYNFTD 431

Qy        396 STVLQL--TKKVNNIETGWALGATFHL 420
```

```
                |:  |:    |: :   || ||   :|
Db        432 SSWEQIHFMGKIKDSNAGWTLGYMLNL 458


RESULT 12
YY4E_CAEEL
ID   YY4E_CAEEL     STANDARD;      PRT;   485 AA.
AC   Q18411;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical 54.3 kDa protein C33H5.14 in chromosome IV.
GN   C33H5.14.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Bradshaw H., Stellyes L.;
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U41007; AAA82272.1; -.
DR   PIR; T34147; T34147.
DR   WormPep; C33H5.14; CE04157.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hypothetical protein; Transmembrane; Hydrolase.
FT   TRANSMEM    439    459       POTENTIAL.
SQ   SEQUENCE   485 AA;  54309 MW;  DE64D1ADC20F581E CRC64;

  Query Match             17.0%;  Score 382.5;  DB 1;  Length 485;
  Best Local Similarity   27.5%;  Pred. No. 3.9e-23;
  Matches  119;  Conservative   84;  Mismatches  153;  Indels   77;  Gaps   18;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDS------VKPGLSAFVDQPKQGAE 102
              ||:: ||||:|||: ||| :: : | |  ::  : :|       | ||||:| |:|:|  |
Db         25 YGVICDAGSSGTRLFVYT-LKPLSGGLTNIDTLIHESEPVVKKVTPGLSSFGDKPEQVVE 83

Qy        103 TVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFL-VPK 161
               :  ||  |:: ||       :| ::: ||||:||||| :  |::  ::   :    | |
Db         84 YLTPLLRFAEEHIPYEQLGETDLLIFATAGMRLLPEAQKDAIIKNLQNGLKSVTALRVSD 143

Qy        162 GSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEK------ 215
              :: |: |  |||::|: ||:: |:         || :|:|||| || |    ||
Db        144 SNIRIIDGAWEGIYSWIAVNYILGRFDKENDSKVGMIDMGGASVQIAFEIANEKESYNGG 203

Qy        216 TLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRSACLPR 275
              : :   | : : | :   ||:|: ::||:|           :|   |        :| ||
Db        204 NVYEINLGSIETNEDYK--YKIYSTTFLGYGANEGLKKYENSLVKSGNS----NDSCSPR 257

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPE---------------- 319
              |    : |    ::  |  ||  :: | |:|  :: |   ||
Db        258 GLNR--LIG----EFTVNGTGE--WDVCLAQVSSLI-GDKAQPSCPNPTCFLRNVIAPSV 308

Qy        320 EVQRGSFYAFS-YYYDRAVDTDMIDYEKGGILKVEDFERKAREVC----DNLENFTSGSP 374
              :     | || |:|  :  '     ::  ||    : |  : |: |    :::::    :
Db        309 NLSTVQLYGFSEYWYTTS------NFGSGGEYHYQKFTDEVRKYCQKDWNDIQDGFKRNE 362

Qy        375 F----------LCMDLSYITALLKDGFGFADST--VLQLTKKVNNIETGWALGA------ 416
              |          |   :::|::| |||   | |  : |   |:   |  |||||
Db        363 FPNADIERLGTNCFKAAWVTSVLHDGFN-VDKTKHLFQSVLKIAGEEMQWALGAMLYHSK 421
```

```
Qy        417 --TFHLLQSLGIS 427
                |:||: | ::
Db        422 DLKFNLLEQLEVA 434


RESULT 13
ENP1_BOVIN
ID   ENP1_BOVIN     STANDARD;      PRT;   513 AA.
AC   O18956;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE   (NTPDase1) (Ecto-ATP diphosphohydrolase) (ATPDase) (Lymphoid cell
DE   activation antigen) (CD39 antigen) (Ecto-apyrase).
GN   ENTPD1 OR CD39.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Aortic endothelium;
RA   Chang A.S., Garcia R.L., Chang S.M., Schilling W.P.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 97-103; 123-133; 136-140; 145-168 AND 459-471.
RC   TISSUE=Aorta;
RX   MEDLINE=97115858; PubMed=8955160;
RA   Kaczmarek E., Koziak K., Sevigny J., Siegel J.B., Anrather J.,
RA   Beaudoin A.R., Bach F.H., Robson S.C.;
RT   "Identification and characterization of CD39/vascular ATP
RT   diphosphohydrolase.";
RL   J. Biol. Chem. 271:33116-33122(1996).
CC   -!- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC       nucleotides to regulate purinergic neurotransmission. Could also
CC       be implicated in the prevention of platelet aggregation.
CC       Hydrolyzes ATP and ADP equally well.
CC   -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.
CC   -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF005940; AAB62382.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT   DOMAIN        1     16       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     17     37       POTENTIAL.
FT   DOMAIN       38    481       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    482    502       POTENTIAL.
FT   DOMAIN      503    513       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     73     73       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    227    227       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    245    245       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    307    307       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    336    336       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    373    373       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    460    460       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     97     97       K -> N (IN REF. 2).
FT   CONFLICT    101    103       INV -> CGF (IN REF. 2).
FT   CONFLICT    464    464       K -> V (IN REF. 2).
SQ   SEQUENCE   513 AA;  58113 MW;  20FE98F27B6D2F96 CRC64;

  Query Match             16.7%;  Score 375;  DB 1;  Length 513;
  Best Local Similarity   26.7%;  Pred. No. 1.7e-22;
  Matches  130;  Conservative   72;  Mismatches  177;  Indels  108;  Gaps   19;

Qy          8 VFFMLVVSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYTF 67
              :  :|  ||: : ::          |:  :   | ||     :||: ||||: | :::| :
Db         17 ILSILGFSCIIAVIA-----LLALGLTQNKALPENVK----FGIVLDAGSSHTSLYIYRW 67

Qy         68 VQKMPGQLPILEGEVFDSVK-PGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTPVV 126
                :      ::      :|| ||:| |  :  :    :   :| |:  ||     :|||
Db         68 PAEKENDTGVVTQIEESNVKGPGISGFAKKVNEINVYLTACMERAQKVIPSIQHMETPVY 127

Qy        127 LKATAGLRLL-PEHK--AKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTVNFL 183
              | ||||:||| |:|  |  :|  |     : ||        |::||:|| : |:|||:|
Db        128 LGATAGMRLLRMENKQMADKILAAVASSISEYPF--DFQGARIISGQEEGAYGWITVNYL 185

Qy        184 TGQL-----------HGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMF 231
              |:                 ||| | ||||||||||||:|| | |  ::|   | | ::
Db        186 LGKFTQKLSWFNLKPSKDDTQETYGALDLGGASTQITFVPQNETT--ESPNNNL-YFRLY 242

Qy        232 NSTYKLYTHSYLGFGLKAARLATLGALETEGTDG-------------------------- 265
                | :||||:| :|   | |  | ||  :||:|
Db        243 GKNYSVYTHSFLCYGKDQALLQKL-ALGLQGTNGIIHEPCFHSRYMRKIKMSVLNEGFCT 301

Qy        266 --HTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRV-----------V 311
                |   |:  |           :: :  ||      |: |   ::::
Db        302 KRHELNSSFYP--------LVDIEIRGAGN------FQRCRQSIIQLFNTSYCPYSSCSF 347

Qy        312 RGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAREVC-------- 363
                |    |   | |:| || |       : ::      : ||    |  | |
Db        348 NGVFLPPLHGQFGAFSAFYYV------MEFLNLTSEESVSVEQLTEKLEEFCAQRWEEVQ 401

Qy        364 ----DNLENFTSGSPFLCMDLSYITALLKDGFGF-ADS-TVLQLTKKVNNIETGWALGAT 417
                  : | : |    |   :||  || :|: | |:|   :    || : : || ||
Db        402 KNFGEVKEKYLSE---YCFSGTYILVLLLNGYHFTAESWKNIHFMNKVRSTDVGWTLGYM 458

Qy        418 FHLLQSL 424
                :|   :
Db        459 LNLTNKI 465


RESULT 14
YND1_YEAST
ID   YND1_YEAST     STANDARD;      PRT;   630 AA.
AC   P40009;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Golgi apyrase (EC 3.6.1.5) (ATP-diphosphatase) (Adenosine
DE   diphosphatase) (ADPase) (ATP-diphosphohydrolase) (Golgi nucleoside
DE   diphosphatase).
GN   YND1 OR YER005W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   STRAIN=S288c;
RX   MEDLINE=99340091; PubMed=10409709;
RA   Gao X.D., Kaigorodov V., Jigami Y.;
RT   "YND1, a homologue of GDA1, encodes membrane-bound apyrase required
RT   for Golgi N- and O-glycosylation in Saccharomyces cerevisiae.";
RL   J. Biol. Chem. 274:21450-21456(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   PubMed=9169868;
RA   Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA   Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
```

```
RA   Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA   Hunicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA   Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA   Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA   Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
RL   Nature 387:78-81(1997).
CC   -!- FUNCTION: CATALYZES THE HYDROLYSIS OF PHOSPHOANHYDRIDE BONDS OF
CC       NUCLEOSIDE TRI-AND DI-PHOSPHATES. HAS EQUAL HIGH ACTIVITY TOWARD
CC       ADP/ATP, GDP/GTP, AND UDP/UTP AND APPROXIMATELY 50% LESS TOWARD
CC       CDP/CTP AND THIAMINE PYROPHOSPHATE. HAS NO ACTIVITY TOWARD GMP.
CC       REQUIRED FOR GOLGI GLYCOSYLATION AND CELL WALL INTEGRITY.
CC   -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC   -!- PATHWAY: Glycosylation.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE BOUND; GOLGI.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF203695; AAF17573.1; -.
DR   EMBL; U18778; AAB64538.1; -.
DR   PIR; S50463; S50463.
DR   SGD; S0000807; YND1.
DR   GO; GO:0006486; P:protein amino acid glycosylation; IMP.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Transmembrane; Golgi stack.
FT   DOMAIN        1    500       LUMENAL (POTENTIAL).
FT   TRANSMEM    501    517       POTENTIAL.
FT   DOMAIN      518    630       CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   630 AA;  71851 MW;  02F8D24A78212544 CRC64;

  Query Match           16.5%;  Score 371;  DB 1;  Length 630;
  Best Local Similarity  26.0%;  Pred. No. 4.6e-22;
  Matches  116;  Conservative   75;  Mismatches  164;  Indels   92;  Gaps   17;

Qy         49 YGIMFDAGSTGTRIHVYTF------------VQKMPGQLPILEGE---VFDSVKPGLSA 92
              :||: |||||:|:|||||: :               | :   :| :  |    |  : ||||:
Db         10 FGIVIDAGSSGSRIHVFKWQDTESLLHATNQDSQSILQSVPHIHQEKDWTF-KLNPGLSS 68

Qy         93 FVDQPKQGAET-VQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEV-KE 150
              |  :|:   :: :: ||: ||: || |||   || ::||||:||||:    ::|  : :
Db         69 FEKKPQDAYKSHIKPLLDFAKNIIPESHWSSCPVFIQATAGMRLLPQDIQSSILDGLCQG 128

Qy        151 IFRKSPFLVP--KGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQE-----TVGTLDLGGA 203
              :   : |||      : :: |: ||:: |: :|:| |  : :  |      | | :|:|||
Db        129 LKHPAEFLVEDCSAQIQVIDGETEGLYGWLGLNYLYGHFNDYNPEVSDHFTFGFMDMGGA 188

Qy        204 STQITFLPQFEKTLE-----------QTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARL 252
              |||| | |      :             ::  | |  :::| ||        :||||   ||
Db        189 STQIAFAPHDSGEIARHRDDIATIFLRSVNGDLQKWDVFVST-------WLGFGANQARR 241

Qy        253 ATLGAL-----------ETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGFE 301
                | |           | :           |:||    :: |    :   |:   |   :
Db        242 RYLAQLINTLPENTNDYENDDFSTRNLNDPCMPRGSSTDFEFKDTIFHIAGSGNYEQCTK 301

Qy        302 PCYAEVLR--------VVRGKLHQPE-EVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKV 352
                | :|:          :  :| | :      |   | |:  | |     :: ||
Db        302 SIYPLLLKNMPCDDEPCLFNGVHAPRIDFANDKFIGTSEYWYTANDV---FKLGGEYNF 357

Qy        353 EDFERKAREVCDN-----LENFTSG------SPFL---CMDLSYITALLKDGFGFA---- 394
              : | :  || |::     | |   |         ||   |   :::  :| :||
Db        358 DKFSKSLREFCNSNWTQILANSDKGVYNSIPENFLKDACFKGNWVLNILHEGFDMPRIDV 417

Qy        395 ------DSTVLQLTKKVNNIETGWALG 415
                   |  : |  :||   |  | ||
```

```
Db        418 DAENVNDRPLFQSVEKVEERELSWTLG 444


RESULT 15
ENP1_PIG
ID   ENP1_PIG       STANDARD;      PRT;   510 AA.
AC   Q9MYU4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE   (NTPDase1) (Ecto-ATP diphosphohydrolase) (ATPDase) (Lymphoid cell
DE   activation antigen) (Ecto-apyrase) (CD39 antigen).
GN   ENTPD1 OR CD39.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Aortic endothelium;
RX   MEDLINE=20325090; PubMed=10866813;
RA   Lemmens R., Vanduffel L., Kittel A., Beaudoin A.R., Benrezzak O.,
RA   Sevigny J.;
RT   "Distribution, cloning, and characterization of porcine nucleoside
RT   triphosphate diphosphohydrolase-1.";
RL   Eur. J. Biochem. 267:4106-4114(2000).
RN   [2]
RP   SEQUENCE OF 202-220.
RC   TISSUE=Pancreas;
RX   MEDLINE=97115858; PubMed=8955160;
RA   Kaczmarek E., Koziak K., Sevigny J., Siegel J.B., Anrather J.,
RA   Beaudoin A.R., Bach F.H., Robson S.C.;
RT   "Identification and characterization of CD39/vascular ATP
RT   diphosphohydrolase.";
RL   J. Biol. Chem. 271:33116-33122(1996).
CC   -!- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC       nucleotides to regulate purinergic neurotransmission. Could also
CC       be implicated in the prevention of platelet aggregation.
CC       Hydrolyzes ATP and ADP equally well.
CC   -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC   -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC   -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION FOUND IN VASCULAR
CC       ENDOTHELIUM, SMOOTH MUSCLE, SPLEEN AND LUNG.
CC   -!- PTM: CLEAVED IN TWO POLYPEPTIDES THAT SEEM TO STAY TOGETHER BY
CC       NONCOVALENT INTERACTIONS.
CC   -!- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ133746; CAB95871.1; -.
DR   InterPro; IPR000407; GDA1_CD39_NTPase.
DR   Pfam; PF01150; GDA1_CD39; 1.
DR   PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW   Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT   CHAIN         1    510       ECTONUCLEOSIDE TRIPHOSPHATE
FT                                DIPHOSPHOHYDROLASE 1.
FT   CHAIN         1    201       ECTONUCLEOSIDE TRIPHOSPHATE
FT                                DIPHOSPHOHYDROLASE 1, 27 KDA SUBUNIT.
FT   CHAIN       202    510       ECTONUCLEOSIDE TRIPHOSPHATE
FT                                DIPHOSPHOHYDROLASE 1, 54 KDA SUBUNIT.
FT   DOMAIN        1     16       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     17     37       POTENTIAL.
FT   DOMAIN       38    477       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    478    498       POTENTIAL.
```

```
FT   DOMAIN      499    510       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     73     73       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    245    245       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    274    274       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    291    291       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    370    370       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    457    457       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    203    203       G -> S (IN REF. 2).
SQ   SEQUENCE   510 AA;  57757 MW;  82F86869040D7735 CRC64;

  Query Match             16.3%;  Score 367;  DB 1;  Length 510;
  Best Local Similarity   27.7%;  Pred. No. 7.2e-22;
  Matches  125;  Conservative   68;  Mismatches  167;  Indels   92;  Gaps   18;

Qy         32 GIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVK-PGL 90
              |:  :   | ||     :||: ||||: | :::| :  :       ::       :| ||:
Db         36 GLTQNKPLPENVK----FGIVLDAGSSHTSLYIYKWPAEKENDTGVVSQVEECKLKGPGI 91

Qy         91 SAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLL---PEHKAKALLFE 147
              | |  :  :    ::  :| |:  :|:|   :||| | ||||:|||    |: |  :|
Db         92 SEFAKKLGEIDIYLEACMERARTVVPKSQHAETPVYLGATAGMRLLRMKNENLASKILST 151

Qy        148 VKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTVNFLTGQL-----------HGHRQETV 195
              | |   : ||         |:|||:|| : |:|:|:|  :              |  |||
Db        152 VAESITRYPF--DFQGARIITGQEEGAYGWITINYLLDKFIQKSGWFNLKPRKGDTQETY 209

Qy        196 GTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATL 255
              | |||||||||||:||  : || :|   |  | ::   | :||||:| :|   | |  |
Db        210 GALDLGGASTQITFVPQ-NQVLE-SPENTL-HFRLYGKNYSVYTHSFLCYGKDQALLQKL 266

Qy        256 GALETEGTDG-----------------HTFRSACLPRWLEAEWIFGGVKYQYGGNQEGE 297
             : : |:|                  |-: : |  |:|     |  :: |  |:
Db        267 -TKDLKNTNGTIHEPCFHSGYQRRMNVSHLYEAPCTRRFL-TSLPFPELEIQGTGD---- 320

Qy        298 VGFEPCYAEVLRVVRGKLHQPEE----------VQRGSFYAFSYYYDRAVDTDMIDYEKG 347
             |: |  : :| :   : |              : :| | ||| :|          |  |
Db        321 --FQKC-QQSIRPLFNTSYCPYSRCSFDGVFLPLPQGDFAAFSAFY----------YVMG 367

Qy        348 GILKVEDFERKAREVCDNLENFTSGSPF----------------LCMDLSYITALLKDG 390
             :    :      :|   || | |  |:                |   :||  ||  |
Db        368 FLNLTSEEGSFQSKVTSTLEAFCS-RPWAELQMYFGDVKEKYLSEYCFSGTYILTLLLSG 426

Qy        391 FGFADST--VLQLTKKVNNIETGWALGATFHL 420
             : |   |   :    || :   || ||   :|
Db        427 YHFTAETWKNIHFMGKVQSTSVGWTLGYMLNL 458


Search completed: January 29, 2004, 10:04:10
Job time : 10.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:58:54 ; Search time 16.5 Seconds
(without alignments)
2494.557 Million cell updates/sec

Title: US-10-091-085-7
Perfect score: 2252
Sequence: 1 MATSWGTVFFMLVVSCVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	616.5	27.4	479	2	T23508	hypothetical prote
2	535.5	23.8	556	2	T39109	probable guanosine
3	488.5	21.7	518	2	A40732	guanosine-diphosph
4	484	21.5	455	2	S48859	nucleoside triphos
5	441.5	19.6	454	2	JC4616	apyrase (EC 3.6.1.
6	426.5	18.9	516	2	G84442	probable nucleosid
7	398	17.7	483	2	D86276	hypothetical prote
8	382.5	17.0	485	2	T34147	hypothetical prote
9	374.5	16.6	557	2	T16696	hypothetical prote
10	371	16.5	630	2	S50463	hypothetical prote
11	366	16.3	510	2	I56242	lymphoid cell acti
12	345	15.3	572	2	T40856	probable nucleotid
13	332.5	14.8	405	2	E86276	hypothetical prote
14	279	12.4	1052	2	T04439	hypothetical prote
15	242	10.7	508	2	C86276	7A19.33 protein -
16	148	6.6	628	2	A55421	nucleoside-triphos
17	108.5	4.8	1019	2	T40813	probable cell divi
18	106.5	4.7	535	2	F97910	glucan 1,6-alpha-g
19	105.5	4.7	553	1	G1BPSV	gene 1 protein - s
20	104.5	4.6	535	2	C95040	glucan 1,6-alpha-g
21	100.5	4.5	1151	2	H71347	hypothetical prote
22	99.5	4.4	774	2	T14555	DNA polymerase hom
23	99	4.4	711	2	S66749	hypothetical prote
24	99	4.4	726	2	AB0122	probable ferric si
25	98.5	4.4	4427	2	PN0637	polyketide synthas
26	98	4.4	497	1	JC2192	subtilisin-like pr
27	96.5	4.3	287	2	T40138	probable ribosomal
28	96	4.3	477	2	JS0597	t-plasminogen acti
29	95.5	4.2	301	2	E71482	phosphatidylserine
30	93	4.1	478	2	F90497	hypothetical prote
31	93	4.1	528	2	S13641	H+-transporting tw
32	93	4.1	594	2	C96638	hypothetical prote
33	92.5	4.1	684	2	T47694	probable serine/th
34	92.5	4.1	1252	2	D71810	probable type II D
35	92	4.1	290	2	A87312	coproporphyrinogen
36	92	4.1	424	2	F69476	acetylornithine am
37	92	4.1	929	2	T17392	vrlL protein - Dic
38	92	4.1	1038	2	H90053	hypothetical prote
39	92	4.1	1220	2	T18291	patched protein -
40	92	4.1	1549	2	S50705	hypothetical prote
41	91.5	4.1	331	2	D83480	cytochrome o ubiqu
42	91.5	4.1	391	2	H89777	capsular polysacch
43	91.5	4.1	490	2	F82973	betaine aldehyde d
44	91.5	4.1	604	2	F89453	protein F35H12.4 [
45	91	4.0	417	2	H70174	glycine hydroxymet

ALIGNMENTS

RESULT 1
T23508
hypothetical protein K08H10.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C;Accession: T23508
R;Gardner, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19750
A;Accession: T23508
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-479 <WIL>
A;Cross-references: EMBL:Z83113; PIDN:CAB05544.1; GSPDB:GN00023; CESP:K08H10.4
A;Experimental source: clone K08H10
C;Genetics:
A;Gene: CESP:K08H10.4
A;Map position: 5
A;Introns: 36/3; 83/3; 189/1; 300/2; 412/3
C;Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 27.4%; Score 616.5; DB 2; Length 479;
Best Local Similarity 34.1%; Pred. No. 7.4e-44;
Matches 149; Conservative 82; Mismatches 167; Indels 39; Gaps 16;

```
Qy          9 FFMLVVS--CVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYT 66
              | :|::|   : | |: : |  |  |  : :    |       : |: ||||||||:|:|
Db          6 FSILLISFFSLLSVVTTKTQY-WCHGDGVLN----NQHTCRFFTIVIDAGSTGTRLHLYK 60

Qy         67 FVQK-------MPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSH 119
              |:         || :   :| |:|  ||||||:|    |    |:::: ||: |:  :|
Db         61 FIHDPAIASHGMPFK---VEKEIFQEVKPGLSSFAKSPSSAADSLEPLLQRARKEVPHFM 117

Qy        120 WKKTPVVLKATAGLRLLPEHKAKALLFEVKE-IFRKSPFLVPKGSVSIMTGQDEGIFAWV 178
              |:|||: ||||||||||    |  :|  |:| ||    |      :|::| | |||:::|
Db        118 WEKTPITLKATAGLRLLPGDMADDILESVEERIFNSGFFAAFPDAVNVMPGSDEGVYSWF 177

Qy        179 TVNFLTGQLH------GHR---QETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFE 229
              |:|  |   |        ||:     :|   |||| |||:|: |  |    :   ||    :
Db        178 TLNILLETLFTDEPTVGHKPAAHRSVAAFDLGGGSTQLTYWPNNEAVFSEHV-GYERDID 236

Qy        230 MFNSTYKLYTHSYLGFGLKAARLATLGALETEG--TDGHTFRSACLPRWLE-AEWIFGGV 286
              |    :|:|||:|| || ||||  |  |||:      |   ::|:|   :  || :  :
Db        237 FFGHHIRLFTHSFLGNGLIAARLNIL-QLETDNEIESTHQLITSCMPEGYQLTEWEY-AL 294

Qy        287 KYQYGGNQEGEVGFEPCYAEVLRVVR-GKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYE 345
              |: :  |       || ||       |   ::    |::    | |||::|||::: ::
Db        295 KF-WNINGSSSHSFESCYGTTKNFVESSEIMHLRELKGSPVYLFSYFFDRALNSGLVKGN 353

Qy        346 KGGILKVEDFERKAREVC----DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQL 401
              :|| :::  |:  |   |       ::: :   |: |:||:|| :||:||: | |:  | |
Db        354 EGGKIELRQFKEAAEIACRREKTEIDDGSHWMPWQCLDLTYIYSLLRDGYQFEDNQPLVL 413
```

```
Qy        402 TKKVNNIETGWALGATF 418
              ||:  :|  |  |  |
Db        414 AKKIKGMEVSWGQGLAF 430
```

RESULT 2
T39109
probable guanosine-diphosphatase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T39109
R;Barrell, B.G.; Rajandream, M.A.; Quail, M.; Seegar, K.; Harris, D.
submitted to the EMBL Data Library, October 1999
A;Reference number: Z21828
A;Accession: T39109
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-556 <BAR>
A;Cross-references: EMBL:AL121741; PIDN:CAB57338.1; GSPDB:GN00066; SPDB:SPAC824.08
A;Experimental source: strain 972h-; cosmid c824
C;Genetics:
A;Gene: SPDB:SPAC824.08
A;Map position: 1

  Query Match 23.8%; Score 535.5; DB 2; Length 556;
  Best Local Similarity 34.1%; Pred. No. 6.1e-37;
  Matches 143; Conservative 61; Mismatches 158; Indels 57; Gaps 13;

```
Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGLL 108
              | :| ||||||:|:||| |    |  | || | |  ::||||:|   |:  | ::  ||
Db        134 YVLMIDAGSTGSRVHVYQFNNCNPS--PKLEEEFFKMIEPGLSSFAGDPEGAAASLDPLL 191

Qy        109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKS-PFLVPKGSVSIM 167
              : | :::|  : : :|: :||||||||  | :|||:|  |::      || : |  |||:
Db        192 DYAMENVPEEYRRCSPIAVKATAGLRLTGESEAKAILKSVRQHLENDYPFPIVKDGVSIL 251

Qy        168 TGQDEGIFAWVTVNFLTGQLHGH-RQETVGTLDLGGASTQITFLPQFEKTLEQTPRG--- 223
              |  |||:||:|:|:| | | |     ||  :||||||||: | |:|    |  |
Db        252 EGSMEGIYAWITINYLLGTLGGKATHSTVAVMDLGGASTQLVFEPRFASDGESLVDGDHK 311

Qy        224 YLTSFEMFNSTYKLYTHSYLGFGLKAARL----------ATLGALETEGTDGHTFRSAC 272
              |:  :      |:|| ||:||:||| ||           |   :||  | |  :    |
Db        312 YVLDYN--GEQYELYQHSHLGYGLKEARKLIHKFVLNNAEALKESLELLG-DSTSIIHPC 368

Qy        273 L----------PRWLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVR----GKLHQP 318
              |          :  :| :| |   :   |  : :  | :    ||      :|||
Db        369 LHLNASLTHPDSKSEASEVVFVGPSLAHLSLQCRGIAEKALYKDKNCPVRPCSFNGVHQP 428

Qy        319 ---EEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAREVC------------ 363
                 |      |  ||:|||     ||         :|| :  |  ||
Db        429 KFTETFTDSPIYLISYFYDR-----MISLGMPSTFTIEDMKYLANSVCSGPTYWQDAFSL 483

Qy        364 -DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATFHLL 421
               | |:      |  |:||:|: :||  |:   ::  |   ||::| | || |||:  :|
Db        484 TDALKELKE-EPEWCLDLNYMISLLSVGYEIPNNRQLHTAKKIDNKELGWCLGASLSML 541
```

RESULT 3
A40732
guanosine-diphosphatase (EC 3.6.1.42) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YEL042w
C;Species: Saccharomyces cerevisiae
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 02-Sep-2000
C;Accession: A40732; B40732; S30837; S50502
R;Abeijon, C.; Yanagisawa, K.; Mandon, E.C.; Hausler, A.; Moremen, K.; Hirschberg, C.B.;
J. Cell Biol. 122, 307-323, 1993
A;Title: Guanosine diphosphatase is required for protein and sphingolipid glycosylation
A;Reference number: A40732; MUID:93308137; PMID:8391537
A;Accession: A40732
A;Molecule type: DNA
A;Residues: 1-518 <ABE>
A;Cross-references: EMBL:L19560; NID:g349392; PIDN:AAA34656.1; PID:g349393
A;Note: sequence extracted from NCBI backbone (NCBIN:134708, NCBIP:134711)
A;Accession: B40732
A;Molecule type: protein
A;Residues: 125-144;238-257;276-281;366-374;399-412 <AB2>
R;Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor, I.
submitted to the EMBL Data Library, February 1993
A;Reference number: S30812
A;Accession: S30837
A;Molecule type: DNA
A;Residues: 1-518 <MUL>
A;Cross-references: GB:U18779; EMBL:L10830; NID:g603625; PIDN:AAB65000.1; PID:g603637
R;Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A;Description: The sequence of S. cerevisiae cosmids 8199, 8334, and 9871.
A;Reference number: S50491
A;Accession: S50502
A;Molecule type: DNA
A;Residues: 1-518 <DIE>
A;Cross-references: EMBL:U18779; NID:g603625; PIDN:AAB65000.1; PID:g603637; MIPS:YEL042w
C;Genetics:
A;Gene: SGD:GDA1
A;Cross-references: SGD:S0000768; MIPS:YEL042w
A;Map position: 5L
C;Function:
A;Description: hydrolase
C;Superfamily: nucleoside triphosphatase chromatin-associated
C;Keywords: glycoprotein; Golgi apparatus; hydrolase; transmembrane protein
F;10-24/Domain: transmembrane #status predicted <TMM>
F;41,280,335/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match 21.7%; Score 488.5; DB 2; Length 518;
  Best Local Similarity 33.0%; Pred. No. 5e-33;
  Matches 143; Conservative 56; Mismatches 159; Indels 75; Gaps 15;

```
Qy         49 YGIMFDAGSTGTRIHVYTF-VQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGL 107
              | || ||||||:|:|:| | |    |  | |  | || ::||||:|       | ::  |
Db         93 YVIMIDAGSTGSRVHIYKFDVCTSP---PTLLDEKFDMLEPGLSSFDTDSVGAANSLDPL 149

Qy        108 LEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKS-PFLVPKG-SVS 165
              |:|| : :|      ||| :||||||||| : |:  :|  |::    |  || | :|  ||
Db        150 LKVAMNYVPIKARSCTPVAVKATAGLRLLGDAKSSKILSAVRDHLEKDYPFPVVEGDGVS 209

Qy        166 IMTGQDEGIFAWVTVNFLTGQL--HGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRG 223
              || | :||:|||:| |:| | :  :| :   |  |||| |||| | | |    :   .
Db        210 IMGGDEEGVFAWITTNYLLGNIGANGPKLPTAAVFDLGGGSTQIVFEPTFPINEKMVDGE 269

Qy        224 YLTSFEMFNSTYKLYTHSYLGFGLKAAR-----------LATLGALETEGTDGHTFRSAC 272
              :    :  :  | ||  |:||:|||  |           |      |: : |  |   | |
Db        270 HKFDLKFGDENYTLYQFSHLGYGLKEGRNKVNSVLVENALKDGKILKGDNTKTHQLSSPC 329

Qy        273 LPRWLEA--EWI-----------FGGVKYQYGG----------NQEGEVGFEPCYAEVLR 309
              ||  : |  | :           | |  |           |:: :    ||
Db        330 LPPKVNATNEKVTLESKETYTIDFIGPDEPSGAQCRFLTDEILNKDAQCQSPPCSF---- 385

Qy        310 VVRGKLHQPEEV----QRGSFYAFSYYYDRAVDTDM-IDYEKGGILKVEDFERKAREVC- 363
                :|||  |    :   | |||:|||    | : :       : :    || ||
Db        386 ---NGVHQPSLVRTFKESNDIYIFSYFYDRTRPLGMPLSF------TLNELNDLARIVCK 436

Qy        364 ------------DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIET 410
                          :|:   | | | |:|||:  :||  |:       |:  ||: | |
Db        437 GEETWNSVFSGIAGSLDELESDSHF-CLDLSFQVSLLHTGYDIPLQRELRTGKKIANKEI 495

Qy        411 GWALGATFHLLQS 423
              || |||:  ||::
Db        496 GWCLGASLPLLKA 508
```

RESULT 4
S48859
nucleoside triphosphatase precursor, chromatin-associated - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 29-Sep-1999
C;Accession: S65147; S48859

R;Hsieh, H.L.; Tong, C.G.; Thomas, C.; Roux, S.J.
Plant Mol. Biol. 30, 135-147, 1996
A;Title: Light-modulated abundance of an mRNA encoding a calmodulin-regulated, chromatin
A;Reference number: S65141; MUID:96197404; PMID:8616230
A;Accession: S65147
A;Molecule type: mRNA
A;Residues: 1-455 <HS2>
A;Cross-references: EMBL:Z32743; NID:g563611; PIDN:CAA83655.1; PID:g563612
C;Superfamily: nucleoside triphosphatase chromatin-associated
C;Keywords: nucleus

Query Match 21.5%; Score 484; DB 2; Length 455;
Best Local Similarity 33.4%; Pred. No. 9.9e-33;
Matches 139; Conservative 62; Mismatches 161; Indels 54; Gaps 14;

```
Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEG-EVFDSVKPGLSAFVDQPKQGAETVQGL 107
              | ::|||||||:||||| | | :   | | :| | :: : ||||:: : |:| |:::  |
Db         44 YAVVFDAGSTGSRIHVYHFNQNL-DLLHIGKGVEYYNKITPGLSSYANNPEQAAKSLIPL 102

Qy        108 LEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIF-RKSPFLVPKGSVSI 166
              || |:| :|    ' |||| | |||||||||     :: :|  |:::   :| | |   :|||
Db        103 LEQAEDVVPDDLQPKTPVRLGATAGLRLLNGDASEKILQSVRDMLSNRSTFNVQPDAVSI 162

Qy        167 MTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPR---- 222
              : |  || : |||||:  | |     :|||:|||| | |: :    :|| :  |:
Db        163 IDGTQEGSYLWVTVNYALGNLGKKYTKTVGVIDLGGGSVQMAYAVS-KKTAKNAPKVADG 221

Qy        223 --GYLTSPEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRS--ACLPRWLE 278
                |:    :    | || |||| || :|:|   |           | ||   ||
Db        222 DDPYIKKVVLKGIPYDLYVHSYLHFGREASRAEILKL---------TPRSPNPCLLAGFN 272

Qy        279 AEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPEEVQRGSF------------ 326
              : : | :::      |   | |   : : :  ||: |   | :|
Db        273 GIYTYSGEEFKATAYTSG-ANFNKCKNTIRKAL--KLNYPCPYQNCTFGGIWNGGGGNGQ 329

Qy        327 ---YAFSYYYDRAVDTDMIDYEKGG-ILKVEDFERKAREVCD-NLENFTSGSPFL----- 376
              :| | ::    || |:|     ||:  | | ||:| |  | |:  |  |||
Db        330 KNLFASSSFFYLPEDTGMVDASTPNFILRPVDIETKAKEACALNFEDAKSTYPFLDKKNV 389

Qy        377 ----CMDLSYITALLKDGFGFADSTVLQLTKKVNN----IETGWALGATFHLLQSL 424
                |||| |   || ||||      :   |::      :|  | ||     : :|
Db        390 ASYVCMDLIYQYVLLVDGFGLDPLQKITSGKEIEYQDAIVEAAWPLGNAVEAISAL 445
```

RESULT 5
JC4616
apyrase (EC 3.6.1.5) precursor - potato
N;Alternate names: adenylpyrophosphatase; ATP-diphosphohydrolase
C;Species: Solanum tuberosum (potato)
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 28-May-1999
C;Accession: JC4616; PC4147
R;Handa, M.; Guidotti, G.
Biochem. Biophys. Res. Commun. 218, 916-923, 1996
A;Title: Purification and cloning of a soluble ATP-diphosphohydrolase (Apyrase) from pot
A;Reference number: JC4616; MUID:96158985; PMID:8579614
A;Accession: JC4616
A;Molecule type: mRNA
A;Residues: 1-454 <HAN>
A;Cross-references: GB:U58597; NID:g1381632; PIDN:AAB02720.1; PID:g1381633
A;Accession: PC4147
A;Molecule type: protein
A;Residues: 59-95;96-131;132-160;236-253;332-345 <HA2>
A;Experimental source: tubers
A;Note: The authors translated the codon GCA for residue 215 as Gly
C;Comment: This enzyme belongs to a family of E-type ATPases, and it catalyzes the hydr
enzyme has nucleotide substrate specificity, divalent cation requirement, and insensiti
olved in starch synthesis.
C;Genetics:
A;Gene: rrop1
C;Superfamily: nucleoside triphosphatase chromatin-associated
C;Keywords: glycoprotein; hydrolase; transmembrane protein
F;1-30/Domain: signal sequence #status predicted <SIG>
F;8-25/Domain: transmembrane #status predicted <TMM>
F;31-454/Product: ATP-diphosphohydrolase #status predicted <MAT>
F;44-65/Region: actin-heat shock protein 70-hexokinase beta-phosphate binding
F;44-65/Region: nucleotide binding #status predicted
F;192-212/Region: actin-heat shock protein 70-hexokinase gamma-phosphate binding
F;192-212/Region: nucleotide binding #status predicted
F;390-410,427-446/Region: hydrophobic carboxyl end
F;151,262/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.6%; Score 441.5; DB 2; Length 454;
Best Local Similarity 28.6%; Pred. No. 3.8e-29;
Matches 130; Conservative 78; Mismatches 178; Indels 69; Gaps 16;

```
Qy         25 NQQTWFEGIFLS-------SMCPINVSA-----------STLYGIMFDAGSTGTRIHVY 65
              || : |   | |:        |:  ||:|           |  | ::|||||||:|:||:
Db          3 NQNSHFIFIILAIFLVLPLSLLSKNVNAQIPLRRHLLSHESEHYAVIFDAGSTGSRVHVF 62

Qy         66 TFVQKMPGQLPILEG-EVFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTP 124
                | :|: | |||     | | : :||||:: : ||  | ::: ||: |:  :|:    :||
Db         63 RFDEKL-GLLPIGNNIEYFMATEPGLSSYAEDPKAAANSLEPLLDGAEGVVPQELQSETP 121

Qy        125 VVLKATAGLRLLPEHKAKALLFEVKEIFR-KSPFLVPKGSVSIMTGQDEGIFAWVTVNFL 183
              : | |||||||:|     |: :|  |: : : :| |       |:|: |  || : |  :|:|
Db        122 LELGATAGLRMLKGDAAEKILQAVRNLVKNQSTFHSKDQWVTILDGTQEGSYMWAAINYL 181

Qy        184 TGQLHGHRQETVGTLDLGGASTQITFL---PQFEKTLE-QTPRGYLTSPEMFNSTYKLYT 239
              | |     : |  |:|||| | |: :      || |  : :     |:     : :  | ||
Db        182 LGNLGKDYKSTTATIDLGGGSVQMAYAISNEQFAKAPQNEDGEPYVQQKHLMSKDYNLYV 241

Qy        240 HSYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVG 299
              |||| :|  | |      |  |         : |      :  : :||| |:     ::|
Db        242 HSYLNYGQLAGRAEIFKASRNES-------NPCALEGCDGYYSYGGVDYKVKAPKKGS-- 292

Qy        300 FEPCYAEVLRVVRG--KLHQPEEVQRGSF--------------YAFSYYYDRAVDTDMI 342
                :    |: |   |::    ::  :|                :| |::||      ::
Db        293 ---SWKRCRRLTRHALKINAKCNIEECTFNGVWNGGGGDGQKNIHASSPFYDIGAQVGIV 349

Qy        343 DYE-KGGILKVEDFERKAREVCD-NLENFTS--------GSPFLCMDLSYITALLKDGFG 392
              | :   : |   :  |:  |  |: :  |          |:|||||  |    || ||||
Db        350 DTKFPSALAKPIQYLNAAKVACQTNVADIKSIFPKTQDRNIPYLCMDLIYEYTLLVDGFG 409

Qy        393 F---ADSTVLQLTKKVNN-IETGWALGATFHLLQS 423
                : ||:   :  |  :   |  ||    |: |
Db        410 LNPHKEITVIHDVQYKNYLVGAAWPLGCAIDLVSS 444
```

RESULT 6
G84442
probable nucleoside triphosphatase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 18-Nov-2002
C;Accession: G84442
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84442
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-516 <STO>
A;Cross-references: GB:AE002093; NID:g3461821; PIDN:AAC32915.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g02970
A;Map position: 2
C;Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 18.9%; Score 426.5; DB 2; Length 516;
Best Local Similarity 30.4%; Pred. No. 8.3e-28;
Matches 126; Conservative 63; Mismatches 153; Indels 73; Gaps 13;

```
Qy         42 NVSASTLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILE--GEVFDSVK--PGLSAFVDQP 97
              |:  |  | :: | |||||||||: :  ::    |: |  |  : |:|  ||||||  | |
```

```
Db         61 NLRGSLRYSVVIDGGSTGTRIHVFGY--RIESGKPVFEFRGANYASLKLHPGLSAFADDP 118

Qy         98 KQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPF 157

Db        119 DGASVSLTELVEFAKGRVPKGMWIETEVRLMATAGMRLLELPVQEKILGVARRVLKSSGF 178

Qy        158 LVPKGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFL------P 211

Db        179 LFRDEWASVISGSDEGVYAWVVANFALGSLGGDPLKTTGIVELGGASAQVTFVSSEPMPP 238

Qy        212 QFEKTLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGAL-------ETEGTD 264

Db        239 EFSRTI---------SFG--NVTYNLYSHSFLHFGQNAAHDKLWGSLLSRDHNSAVEPTR 287

Qy        265 GHTFRSACLPR-------------WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVV 311

Db        288 EKIFTDPCAPKGYNLDANTQKHLSGLLAEESRLSDSFQAGGN------YSQCRSAALTIL 341

Qy        312 RGKLHQPEEVQRGSFYAFSYYYDRAVDTDMI---------DYEKGGILKVEDFERKAREV 362

Db        342 QDGNGRILIIIAGFSFLFFGLGEKAWLSNMISAGERFCGEDWSK---LRVKDPSLHEEDL 398

Qy        363 CDNLENFTSGSPFLCMDLSYITALLKDGFGF-ADSTVLQLTKKVNNIETGWALGA 416

Db        399 LR-----------YCFSSAYIVSLLHDTLGIPLDDERIGYANQAGDIPLDWALGA 442


RESULT 7
D86276
hypothetical protein F7A19.34 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 18-Nov-2002
C;Accession: D86276
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D86276
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-483 <STO>
A;Cross-references: GB:AE005172; NID:g5080801; PIDN:AAD39311.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: nucleoside triphosphatase chromatin-associated

  Query Match           17.7%;  Score 398;  DB 2;  Length 483;
  Best Local Similarity  27.4%;  Pred. No. 1.9e-25;
  Matches  119;  Conservative   73;  Mismatches  161;  Indels   82;  Gaps   17;

Qy         32 GIFLSSMCPINVSASTLYGIMFDAGSTGTRIHV--YTFVQKMPGQLPILEGEVFD----- 84

Db         59 GSLLSRRCKLR------YSVLIDAGSSGTRVHVFGYWFESGKP---------VFDFGEKH 103

Qy         85 ----SVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHK 140

Db        104 YANLKLTPGLSSYADNPEGASVSVTKLVEFAKQRIPKRMFRRSDIRLMATAGMRLLEVPV 163

Qy        141 AKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDL 200

Db        164 QEQILEVTRRVLRSSGFMFRDEWANVISGSDEGIYSWITANYALGSLGTDPLETTGIVEL 223

Qy        201 GGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALET 260

Db        224 GGASAQVTFV-----SSEHVPPEYSRTIAYGNISYTIYSHSFLDYGKDAALKKLLEKLQN 278

Qy        261 EG---TDGHTFRSACLPRWLEAEWIF--GGVKYQYG---------GNQEGEVGFEPCYAE 306

Db        279 SANSTVDG-VVEDPCTPK----GYIYDTNSKNYSSGFLADESKLKGSLQAAGNFSKCRSA 333

Qy        307 VLRVVR-------------GKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKV- 352

Db        334 TFALLKEGKENCLYEHCSIGSTFTPD--LQGSFLATASFYYTAKFFEL--EEKGWLSELI 389

Qy        353 --------EDFERKAREVCDNLENFTSGSPFLCMDLSYITALLKDGFGFA--DSTVLQLT 402

Db        390 PAGKRYCGEEWSKLILEYPTTDEEYLRG---YCFSAAYTISMLHDSLGIALDDESITYAS 446

Qy        403 KK-VNNIETGWALGA 416

Db        447 KAGEKHIPLDWALGA 461


RESULT 8
T34147
hypothetical protein C33H5.14 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Mar-2000
C;Accession: T34147
R;Bradshaw, H.; Stellyes, L.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C33H5.
A;Reference number: Z21482
A;Accession: T34147
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-485 <BRA>
A;Cross-references: EMBL:U41007; PIDN:AAA82272.1; CESP:C33H5.14
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:C33H5.14
A;Introns: 22/1; 83/1; 120/2; 167/1; 269/3; 399/3
C;Superfamily: nucleoside triphosphatase chromatin-associated

  Query Match           17.0%;  Score 382.5;  DB 2;  Length 485;
  Best Local Similarity  27.5%;  Pred. No. 3.9e-24;
  Matches  119;  Conservative   84;  Mismatches  153;  Indels   77;  Gaps   18;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDS------VKPGLSAFVDQPKQGAE 102

Db         25 YGVICDAGSSGTRLFVYT-LKPLSGGLTNIDTLIHESEPVVKKVTPGLSSFGDKPEQVVE 83

Qy        103 TVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFL-VPK 161

Db         84 YLTPLLRFAEEHIPYEQLGETDLLIFATAGMRLLPEAQKDAIIKNLQNGLKSVTALRVSD 143

Qy        162 GSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEK------ 215

Db        144 SNIRIIDGAWEGIYSWIAVNYILGRFDKENDSKVGMIDMGGASVQIAFEIANEKESYNGG 203

Qy        216 TLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRSACLPR 275

Db        204 NVYEINLGSIETNEDYK--YKIYSTTFLGYGANEGLKKYENSLVKSGNS----NDSCSPR 257

Qy        276 WLEAEWIFGGVKYQYGGNQEGEVGFEPCYAEVLRVVRGKLHQPE---------------- 319

Db        258 GLNR--LIG----EFTVNGTGE--WDVCLAQVSSLI-GDKAQPSCPNPTCFLRNVIAPSV 308

Qy        320 EVQRGSFYAFS-YYYDRAVDTDMIDYEKGGILKVEDFERKAREVC----DNLENFTSGSP 374

Db        309 NLSTVQLYGFSEYWYTTS------NFGSGGEYHYQKFTDEVRKYCQKDWNDIQDGFKRNE 362

Qy        375 F----------LCMDLSYITALLKDGFGFADST--VLQLTKKVNNIETGWALGA------ 416

Db        363 FPNADIERLGTNCFKAAWVTSVLHDGFN-VDKTKHLFQSVLKIAGEEMQWALGAMLYHSK 421

Qy        417 --TFHLLQSLGIS 427

Db        422 DLKFNLLEQLEVA 434
```

```
RESULT 9
T16696
hypothetical protein R07E4.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T16696
R;Miller, N.
submitted to the EMBL Data Library, October 1995
A;Description: The sequence of C. elegans cosmid R07E4.
A;Reference number: Z18561
A;Accession: T16696
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-557 <MIL>
A;Cross-references: EMBL:U39652; NID:g1049390; PID:g1049394; PIDN:AAA80403.1; CESP:R07E4
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:R07E4.4
A;Introns: 39/1; 67/1; 145/2; 192/1; 244/3; 317/3; 451/3; 488/2

  Query Match             16.6%;  Score 374.5;  DB 2;  Length 557;
  Best Local Similarity   28.3%;  Pred. No. 2.2e-23;
  Matches  126;  Conservative   70;  Mismatches  159;  Indels   91;  Gaps   21;

Qy         49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDS------VKPGLSAFVDQPKQGAE 102
              ||:: |||||||: || ::     :|  :|  ::|:      : |||| |  :| | ||
Db         44 YGVICDAGSTGTRLFVYNWISTSDSELIQIEPVIYDNKPVMKKISPGLSTFGTKPAQAAE 103

Qy        103 TVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEH-----KAKALLFEVKEIFRKSPF 157
               :: |:|:|:  ||      ||| : ||||:||:|:      : :|:|   | :  | |
Db        104 YLRPLMELAERHIPEEKRPYTPVFIFATAGMRLIPDEYVLIGQKEAVL---KNLRNKLPK 160

Qy        158 L----VPKGSVSIMTGQDEGIFAWVTVNFLTGQLH-----------GH-RQETVGTLDL 200
              :    | |  : |: |: |||::|: ||:  |: :              | ||:||| :|:
Db        161 ITSMQVLKEHIRIIEGKWEGIYSWIAVNYALGKFNKTATLDFPGTSPAHARQKTVGMIDM 220

Qy        201 GGASTQITF-LPQFEK----TLEQTPRGYLTSFEMFNSTYKLYTHSYLGF----GLKAAR 251
              |||| || | ||  :     :|    |      :|   |||:  ::||:    |::
Db        221 GGASAQIAFELPDTDSFSSINVENINLGCREDDSLFK--YKLFVTTFLGYGVNEGIRKYE 278

Qy        252 LATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGN--QEG------------- 296
                 |  |: :  :|   :. |:|  |         |  : | |  : |
Db        279 HMLLSKLKDQ--NGTVIQDDCMPLNLHKT-----VTLENGENFVRRGTGNWNTCSNEVKK 331

Qy        297 ----EVGFEPCYAEVLRVVRGKLHQPE-EVQRGSFYAFSYYYDRAVDTDMIDYEKGGILK 351
                  |   | | ||  :   | :  |   :      | || |:    |.       ||
Db        332 LLNPESSSEVCKAEAAKCYFGAVPAPSIPLSNIEMYGFSEYWYSTHDV----LGLGGQYD 387

Qy        352 VEDFERKAREVCD------NLENFTSGSP--------FLCMDLSYITALLKDGFGFADST 397
               |:  :| :: |          |:     |          |   ::||::| |||   | |
Db        388 AENIAKKTQQYCSKRWSTIQAESKKQLYPRADEERLRTQCFKSAWITSVLHDGFS-VDKT 446

Qy        398 --VLQLTKKVNNIETGWALGA-TFHL 420
                   |     :   |  |||||  :|:
Db        447 HNKFQSVSTIAGQEVQWALGAMIYHM 472


RESULT 10
S50463
hypothetical protein YER005w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 19-Apr-2002
C;Accession: S50463
R;Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A;Description: The sequence of S. cerevisiae cosmids 9537, 9581, 9495, 9867, and lambda
A;Reference number: S50433
A;Accession: S50463
A;Molecule type: DNA
A;Residues: 1-630 <DIE>
A;Cross-references: EMBL:U18778; NID:g603592; PID:g603597; GSPDB:GN00005; MIPS:YER005w
C;Genetics:
A;Gene: SGD:YND1; MIPS:YER005w
A;Cross-references: SGD:S0000807
A;Map position: 5R

  Query Match             16.5%;  Score 371;  DB 2;  Length 630;
  Best Local Similarity   26.0%;  Pred. No. 5.3e-23;
  Matches  116;  Conservative   75;  Mismatches  164;  Indels   92;  Gaps   17;

Qy         49 YGIMFDAGSTGTRIHVYTF------------VQKMPGQLPILEGE---VFDSVKPGLSA 92
              :||: ||||:|:|||||: :              | :   :| :   |    | : ||||:
Db         10 FGIVIDAGSSGSRIHVFKWQDTESLLHATNQDSQSILQSVPHIHQEKDWTF-KLNPGLSS 68

Qy         93 FVDQPKQGAET-VQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEV-KE 150
              |  :|:   :: :: ||: ||: || |||   || ::||||:||||:    ::|  : :
Db         69 FEKKPQDAYKSHIKPLLDFAKNIIPESHWSSCPVFIQATAGMRLLPQDIQSSILDGLCQG 128

Qy        151 IFRKSPFLVP--KGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQE-----TVGTLDLGGA 203
              :   : |||      : :: |: ||:: |: :|:| |  : :  |      | | :|:|||
Db        129 LKHPAEFLVEDCSAQIQVIDGETEGLYGWLGLNYLYGHFNDYNPEVSDHFTFGFMDMGGA 188

Qy        204 STQITFLPQFEKTLE-----------QTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARL 252
              |||| | |     :            ::  | |  :::| ||        :||||   ||
Db        189 STQIAFAPHDSGEIARHRDDIATIFLRSVNGDLQKWDVFVST-------WLGFGANQARR 241

Qy        253 ATLGAL-----------ETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGFE 301
                 | |           | :           |:||    :: |    :   |:   |   :
Db        242 RYLAQLINTLPENTNDYENDDFSTRNLNDPCMPRGSSTDFEFKDTIFHIAGSGNYEQCTK 301

Qy        302 PCYAEVLR--------VVRGKLHQPE-EVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKV 352
                 |  :|:          :  :| | :      |   | |:  | |     :: ||
Db        302 SIYPLLLKNMPCDDEPCLFNGVHAPRIDFANDKFIGTSEYWYTANDV----FKLGGEYNF 357

Qy        353 EDFERKAREVCDN-----LENFTSG------SPFL---CMDLSYITALLKDGFGFA---- 394
              : | :  || |::     | |   |        ||   |   :::  :| :||
Db        358 DKFSKSLREFCNSNWTQILANSDKGVYNSIPENFLKDACFKGNWVLNILHEGFDMPRIDV 417

Qy        395 ------DSTVLQLTKKVNNIETGWALG 415
                     |  : |  :||   |  | ||
Db        418 DAENVNDRPLFQSVEKVEERELSWTLG 444


RESULT 11
I56242
lymphoid cell activation antigen - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Nov-2002
C;Accession: I56242
R;Maliszewski, C.R.; Delespesse, G.J.; Schoenborn, M.A.; Armitage, R.J.; Fanslow, W.C.; M
J. Immunol. 153, 3574-3583, 1994
A;Title: The CD39 lymphoid cell activation antigen. Molecular cloning and structural cha
A;Reference number: I56242; MUID:95015846; PMID:7930580
A;Accession: I56242
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-510 <RES>
A;Cross-references: GB:S73813; NID:g765255; PIDN:AAB32152.1; PID:g765256
C;Superfamily: nucleoside triphosphatase chromatin-associated

  Query Match             16.3%;  Score 366;  DB 2;  Length 510;
  Best Local Similarity   28.2%;  Pred. No. 1e-22;
  Matches  125;  Conservative   72;  Mismatches  172;  Indels   74;  Gaps   19;

Qy         32 GIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVK-PGL 90
              |:  :   | ||     |||: |||| | :::| :  :       ::       || ||:
Db         36 GLTQNKALPENVK----YGIVLDAGSSHTSLYIYKWPAEKENDTGVVHQVEECRVKGPGI 91

Qy         91 SAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLL---PEHKAKALLFE 147
              | || :  :    :  :| |:: ||||  ::||| | ||||:|||     |  |  :|
Db         92 SKFVQKVNEIGIYLTDCMERAREVIPRSQHQETPVYLGATAGMRLLRMESEELADRVLDV 151

Qy        148 VKEIFRKSPFLVPKGSVSIMTGQDEGIFAWVTVNFLTGQL-----------HGHRQETV 195
              |:    ||        |:|||:|| : |:|:|:| |:                : |||
```

```
Db        152 VERSLSNYPF--DFQGARIITGQEEGAYGWITINYLLGKFSQKTRWFSIVPYETNNQETF 209

Qy        196 GTLDLGGASTQITFLPQFEKTLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATL 255

Db        210 GALDLGGASTQVTFVPQ-NQTIE-SPDNAL-QFRLYGKDYNVYTHSFLCYGKDQALWQKL 266

Qy        256 GALETEGTDGHTFRSACL------------------PRWLEAEWIFGGVKYQYGGNQEGE 297

Db        267 -AKDIQVASNEILRDPCFHPGYKKVVNVSDLYKTPCTKRFEMTLPFQQFEIQGIGN---- 321

Qy        298 VGFEPCYAEVLRV------------VRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYE 345

Db        322 --YQQCHQSILELFNTSYCPYSQCAFNGIFLPPLQGDFGAFSAF-YFVMKFLN---LTSE 375

Qy        346 KGGILKVEDFERK-AREVCDNLENFTSG--SPFL---CMDLSYITALLKDGFGF-ADS-T 397

Db        376 KVSQEKVTEMMKKFCAQPWEEIKTSYAGVKEKYLSEYCFSGTYILSLLLQGYHFTADSWE 435

Qy        398 VLQLTKKVNNIETGWALGATFHL 420

Db        436 HIHFIGKIQGSDAGWTLGYMLNL 458


RESULT 12
T40856
probable nucleotide phosphatase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T40856
R;Ramsperger, U.; Pohl, T.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1999
A;Reference number: Z21952
A;Accession: T40856
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-572 <RAM>
A;Cross-references: EMBL:AL121783; PIDN:CAB57847.1; GSPDB:GN00068; SPDB:SPCC11E10.05c
A;Experimental source: strain 972h-; cosmid c11E10
C;Genetics:
A;Gene: SPDB:SPCC11E10.05c
A;Map position: 3

  Query Match             15.3%;  Score 345;  DB 2;  Length 572;
  Best Local Similarity   25.3%;  Pred. No. 7.1e-21;
  Matches  114;  Conservative   69;  Mismatches  156;  Indels  112;  Gaps   16;

Qy         49 YGIMFDAGSTGTRIHVYTF-------VQKMPGQLPILEGEVFD------SVKPGLSAFVD 95

Db          5 YGIFIDAGSSGSRLLIYSWDYDTDSSLSDKVKKLPLIETGIGDGGKWSLKVQPGISSFAN 64

Qy         96 QPKQ-GAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLL-PEHKAKALLFEVKEIFR 153

Db         65 NPKHVGKKHLKELLDFAAHAIPKDVHKETPVFLSATAGMRLLGVDAQNKILSHACRYIKK 124

Qy        154 KSPFLVP--KGSVSIMTGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLP 211

Db        125 NYDFDIPNCSNSIRVIDGKAEGMYGWLATNYLLKTLEEKDTSTVGFLDMGGASVQIAF-- 182

Qy        212 QFEKTLEQTPRGYLTSFEMFNST------------YKLYTHSYLGFGLKAARLATLGAL- 258

Db        183 -------ELPPSQLKNYKDSISTVHIGLQNGQQLEYPLFVTTWLGFGANEAYRRYLGLLI 235

Qy        259 ETE-GTDGHTFRSACLPRWLEAEWIFGGVKYQYGG------------NQEGEVGFEPCY 304

Db        236 ESENGKVGNTLSDPCSLR--GRTYDIDGIEFAGTGDLKQCLKLTYNLLNKDKPCSMDPCN 293

Qy        305 AEVLRVVRGKLHQPEEVQRGSFYAFSYYYDRAVDTDMIDYEKGGILKVEDFERKAREVC- 363

Db        294 FDGISI------PPVDFANTEFVGVSEFWYTTNDV----FDMGGSYHFPNFYKKVDEYCG 343

Qy        364 ----------------------DNLENFTSGSPFLCMDLSYITALLKDGFGFADSTV---- 398

Db        344 TEWETMLSRLYNKELTPSTDENKLEK-------LCFKASWALNVLHEGFDVPKSNTSSND 396

Qy        399 --------------LQLTKKVNNIETGWALG 415

Db        397 AKDGLSVIPAYHSPFTSLEKIERTEVSWTLG 427


RESULT 13
E86276
hypothetical protein F14L17.1 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C;Accession: E86276
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E86276
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-405 <STO>
A;Cross-references: GB:AE005172; NID:g7262666; PIDN:AAF43924.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

  Query Match             14.8%;  Score 332.5;  DB 2;  Length 405;
  Best Local Similarity   26.3%;  Pred. No. 4.9e-20;
  Matches  110;  Conservative   64;  Mismatches  157;  Indels   87;  Gaps   15;

Qy         48 LYGIMFDAGSTGTRIHVYTFVQKMPGQLPILEGEVFDSVKPGLSAFVDQPKQGAETVQGL 107

Db          3 VFGYWFESGKP-----VFDFGEEHYASL---------KLSPGLSSYADNPEGASVSVTKL 48

Qy        108 LEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIM 167

Db         49 VEFAKGRIPKGKLKKSDIRLMATAGMRLLDVPVQEQILDVTRRVLRSSGFKFQDEWATVI 108

Qy        168 TGQDEGIFAWVTVNFLTGQLHGHRQETVGTLDLGGASTQITFLPQFEKTLEQTPRGYLTS 227

Db        109 SGTDEGIYAWVVANHALGSLGGDPLKTTGIVELGGASAQVTFVPS-----EHVPPEFSRT 163

Qy        228 FEMFNSTYKLYTHSYLGFGLKAARLATLGALETE---GTDGHTFRSACLPR--------- 275

Db        164 ISYGNVSYTIYSHSFLDFGQDAAEDKLLESLQNSVAASTGDGIVEDPCTPKGYIYDTHSQ 223

Qy        276 -----WLEAEWIF-GGVKYQYGGNQEGEVGFEPCYAEVLRVVR-GKLHQPEEVQRGSFYA 328

Db        224 KDSSGFLSEESKFKASLQVQAAGD------FTKCRSATLAMLQEGKEN------------ 265

Qy        329 FSYYYDRAVDTDMIDYEKGGILKVEDF----------------------ERKAREVCDNL 366

Db        266 -CAYKHCSIGSTFTPNIQGSFLATENFFHTSKFFGLGEKEWLSEMILAGKRFCGEEWSKL 324

Qy        367 -ENF-TSGSPFL---CMDLSYITALLKDGFGFA-DSTVLQLTKKV--NNIETGWALGA 416

Db        325 KEKYPTTKDKYLHRYCFSSAYIISMLHDSLGVALDDERIKYASKAGKENIPLDWALGA 382


RESULT 14
T04439
hypothetical protein T18B16.150 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
C;Accession: T04439
R;Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansorge, W.; Bancroft, I.; Mewes, H.W.
submitted to the Protein Sequence Database, April 1998
A;Reference number: Z15359
```

```
A;Accession: T04439
A;Molecule type: DNA
A;Residues: 1-1052 <BEV>
A;Cross-references: EMBL:AL021687
A;Experimental source: cultivar Columbia; BAC clone T18B16
C;Genetics:
A;Map position: 4
A;Introns: 79/3; 118/3; 217/1; 295/3; 396/3; 531/2; 815/3
A;Note: T18B16.150

  Query Match             12.4%;  Score 279;  DB 2;  Length 1052;
  Best Local Similarity   25.1%;  Pred. No. 6.2e-15;
  Matches  108;  Conservative   70;  Mismatches  170;  Indels   82;  Gaps   19;

Qy         51 IMFDAGSTGTRIHVY--TFVQKMPGQLPIL----------EGEVFD--SVKPGLSAFV 94
              :::    |||| :||  :   |    |||:             |  :|    :||   |
Db        525 LVYIVSITGTRAYVYQASINYKKDSSLPIVMKSLTEGISRKSRGRAYDRMETEPGFDKLV 584

Qy         95 DQPKQGAET-VQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFR 153
              :  : | :| :: |:: |: ||::  : | : : ||||:| |    :  :|  |  |
Db        585 NN-RTGLKTAIKPLIQWAEKQIPKNAHRTTSLFVYATAGVRRLRPADSSWILGNVWSILA 643

Qy        154 KSPFLVPKGSVSIMTGQDEGIFAWVTVNFLTGQLHG-HRQETVGTLDLGGASTQITFLPQ 212
              ||||   :  | |::| :|  | |  :|: |  |    :: | | |||||:| |:||  :
Db        644 KSPFTCRREWVKIISGTEEAYFGWTALNYQTSMLGALPKKATFGALDLGGSSLQVTFENE 703

Qy        213 FEKTLEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAA----------RLATL------- 255
               |:|  :|      :  : :  : |  :|  |:||  |          :|  :
Db        704 -ERTHNETN----LNLRIGSVNHHLSAYSLAGYGLNDAFDRSVVHLLKKLPNVNKSDLIE 758

Qy        256 GALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGEVGFE------PCYAEVLR 309
              | ||         :  ||      ::|          | ::|: |         | : |
Db        759 GKLE--------MKHPCLNSGYNGQYICSQCASSVQGGKKGKSGVSIKLVGAPNWGECSA 810

Qy        310 VVR-GKLHQPEEVQR--GSFYAFSYYYDRAVDTDMIDYEKGGI---LKVEDFERKAREVC 363
              : :      |:   |  | ||| | ::       : |    :     ::|   | ||| |
Db        811 LAKNAPCALPDGYPRPHGQFYAVSGFF--------VVYRFFNLSAEASLDDVLEKGREFC 862

Qy        364 DN----LENFTSGSPFL---CMDLSYITALLKDGFGFADSTV------LQLTKKVNNIET 410
              |          |  ||:   |    || :||::|   |  :      :  |  | :|:
Db        863 DKAWQVARTSVSPQPFIEQYCFRAPYIVSLLREGLYITDKQIIIGSGSITWTLGVALLES 922

Qy        411 GWALGATFHL 420
              | || :|  |
Db        923 GKALSSTLGL 932


RESULT 15
C86276
7A19.33 protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 18-Nov-2002
C;Accession: C86276
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86276
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-508 <STO>
A;Cross-references: GB:AE005172; NID:g5080800; PIDN:AAD39310.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: nucleoside triphosphatase chromatin-associated

  Query Match             10.7%;  Score 242;  DB 2;  Length 508;
  Best Local Similarity   23.2%;  Pred. No. 2.8e-12;
  Matches  109;  Conservative   78;  Mismatches  161;  Indels  122;  Gaps   21;

Qy         12 LVVSCVCSAV----------SHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
              ::|:||  |:           | ||::                  |  | :: | ||:||
Db         35 VIVACVTIALGLLFIGYSILRSGRNRR----------------VSLHYSVIIDGGSSGT 77

Qy         61 RIHVYTFVQKMPGQLPILE--GEVFDSVK--PGLSAFVDQPKQGAETVQGLLEVAKDSIP 116
              |:||: :  ::    |: :   | : |:|  |||||: | |:  :|:|  |:| ||  :
Db         78 RVHVFGY--RIESGKPVFDFGEENYASLKLSPGLSAYADNPEGVSESVTELVEFAKKRVH 135

Qy        117 RSHWKKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTGQDEGIFA 176
              :  ||: : | ||||:|||     : :|   : : | | |                :||
Db        136 KGKLKKSDIRLMATAGMRLLELPVQEQILDVTRRVLRSSGF----------DFRDE---- 181

Qy        177 WVTVNFLTGQLHGHRQETVGTLDLG------GASTQITFLPQFEKTLEQTPRGYLTSFEM 230
              | :|     ::  : |: :    ||          ::||:     : |  | :  :
Db        182 WASV---ISEILENFQDLMKVYMLGLLLIMRSVRLEVTFV-----STELVPSEFSRTLAY 233

Qy        231 FNSTYKLYTHSYLGFGLKAARLATLGAL---ETEGTDGHTFRSACLPRWLEAEWI----- 282
              | :| ||:||:| ||  ||  ||:     :|        |          |:|:    |
Db        234 GNVSYNLYSHSFLDFGQDAAQEKLSESLYNSAANSTGEGIVPDPCIPKGYILETNLQKDL 293

Qy        283 --FGGVKYQYGGNQEGEVGFEPCYAEVLRVVR------------GKLHQPEEVQRGSFY 327
               |   | ::    :    |  | :    :::                | : |    :|||
Db        294 PGFLADKGKFTATLQAAGNFSECRSAAFAMLQEEKGKCTYKRCSIGSIFTPN--LQGSFL 351

Qy        328 AFSYYY---------DRAVDTDMI---------DYEKGGILKVEDFERKAREVCDNLENF 369
              |   ::          ::   ::||         :: |   |||:    |     :|| :
Db        352 ATENFFHTSKFFGLGEKEWLSEMILAGKRFCGEEWSK---LKVKYPTFKD----ENLLRY 404

Qy        370 TSGSPFLCMDLSYITALLKDGFGFA-DSTVLQLTKKV--NNIETGWALGA 416
                  :|| ::| |  | | |   ::    |    :|   |||||
Db        405 -------CFSSAYIISMLHDSLGVALDDERIKYASKAGEEDIPLDWALGA 447


Search completed: January 29, 2004, 10:06:17
Job time : 17.5 secs
```

THIS PAGE BLANK (USPTO)